Representations of spatial location in language processing

Jens Apel

Ph.D. Thesis

The University of Edinburgh

2010

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Jens Apel)

Abstract

The production or comprehension of linguistic information is often not an isolated task decoupled from the visual environment. Rather, people refer to objects or listen to other people describing objects around them. Previous studies have shown that in such situations people either fixate these objects, often multiple times (Cooper, 1974), or they attend to the objects much longer than is required for mere identification (Meyer, Sleiderink, & Levelt, 1998). Most interestingly, during comprehension people also attend to the location of objects even when those objects were removed (Altmann, 2004). The main focus of this thesis was to investigate the role of the spatial location of objects during language processing.

The first part of the thesis tested whether attention to objects' former locations facilitates language production and comprehension processes (Experiments 1-5). In two initial eye-tracking experiments, participants were instructed to name objects that either changed their positions (Experiment 1) or were withdrawn from the computer screen (Experiment 2) during language production. Production was impaired when speakers did not attend to the original position of the objects. Most interestingly, fixating an empty region in which an object was located resulted in faster articulation and initiation times.

During the language comprehension tasks, participants were instructed to evaluate facts presented by talking heads appearing in different positions on the computer screen. During evaluation, the talking heads changed position (Experiment 3) or were withdrawn from the screen (Experiments 4-5). People showed a strong tendency to gaze at the centre of the screen and only moved towards the head's former locations if the screen was empty and if evaluation was not preceded by an intervening task as tested in Experiment 5. Fixating the former location resulted in faster response time but not in better accuracy of evaluation.

The second part of this thesis investigated the role of spatial location representations in reading (Experiments 6-7). Specifically, I examined to what

extent people reading garden-path sentences regress to specific target words in order to reanalyse the sentences. The results of two eye-tracking experiments showed that readers do not target very precisely. A spatial representation is used, but it appears to be fairly coarse (i.e., only represents whether information is to the left or to the right of fixation).

The findings from this thesis give us a clearer understanding of the influence of spatial location information on language processing. In language production particularly, it appears that spatial location is an integral part of the cognitive model and strongly connected with linguistic and visual representations.

Acknowledgements

First of all, I would like to thank my supervisors Fernanda Ferreira and John Henderson for their support and guidance in writing this thesis. The doors to Fernanda's and John's offices were always open for any kind of problems, questions, and discussions; even about the latest movies.

I would also like to thank Antje Nuthmann who taught me most valuable lessons in Matlab and statistics. Our coffees and lunches were also always proper breaks from work and I am very glad that we became such good friends. Laura Speed, Lisa Kinsella, and Kyleigh Schraeder helped me to transcribe sound files and to run experiments. Robert Maier, Manon Jones, Mateo Obregón, and Paul Engelhardt kindly provided statistical advice. My thanks go also to Bela Usabaev, Manon Jones, Emilia Ellsiepen, and Antje Nuthmann for reading through drafts of this thesis. I also thank the whole psycholinguistic and visual cognition group who helped me in countless little questions. The discussions in our group meetings and reading groups were always very insightful and inspiring.

Furthermore, I want to acknowledge the financial and technical support from the College of Humanities & Social Science of the University of Edinburgh.

I also thank many lovely friends who helped me through this time and who I unfortunately cannot all list here. However, I especially want to thank Bela who showed me to see many things from a different perspective. Furthermore, many thanks to my best and oldest friend Michael. I cannot really express how much it means to me to know you, Pia, and Malte.

My deepest gratitude goes to my parents Klaus und Ulrike Apel and to my sister Yvonne, who have supported me in so many ways in all these years. Vielen, vielen Dank für alles. Ohne Euch, und ohne zu wissen, dass Ihr immer da seid, hätte ich es nie geschafft.

Table of Contents

List of Tables

List of Figures

1 Chapter 1: General Introduction

Imagine the following situation: You are in your local pub and you just met someone new. You introduce yourselves. After a while you go outside for a cigarette. On the way back you try to remember the name of that person but it just slipped your mind. As soon as you reenter the pub and see that person standing at the bar, however, the name just pops into your head.

This example illustrates that information perceived from the external world might influence the retrieval of internal memory. In the current thesis, this effect is studied in greater detail. The main focus of this work was on the influence of spatial location information of objects and their integration in memory. Returning to the anecdote in the pub, I closely investigated the question whether the person needs to be standing at the same spot than he or she stood at the last time? Does the person have to be in the pub in order to facilitate memory or is it sufficient to fixate the spot in which he or she stood the last time?

The psycholinguistic community has investigated the relationship between visual and linguistic information for many decades now and numerous papers have been published in the field of language production, language comprehension, and reading (e.g. Cooper, 1974; Frazier & Rayner, 1982; Henderson & Ferreira, 2004; Kennedy, 1992; Meyer, Sleiderink, & Levelt, 1998; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). However, minimal attention has been focused on memory related issues and how activation of spatial location information influences retrieval of information related to this location.

It has been shown that people fixate objects and words when processing information related to these objects or words (e.g. Cooper, 1974; Frazier & Rayner, 1982; Meyer, et al., 1998; Tanenhaus, et al., 1995). In language production, speakers fixate objects for up to 1500ms before naming them (Griffin, 2001; Meyer, et al., 1998). In spoken language comprehension, listeners gaze at objects when references to these objects appear in speech (Cooper, 1974; Tanenhaus, et al., 1995). In reading, people refixate words during the reanalysis of ambiguous sentences (Frazier & Rayner, 1982). Most crucial for the current study, people gaze at objects even though objects have already been encoded. In language production, in order to identify objects, speakers only need to fixate them for 150ms (Potter, 1975). In language comprehension, people look at a scene, and revisit objects when references of these objects occur in speech. In reading, people return to previously read words, even though they have already been read.

In the current thesis, I investigated the mechanisms that direct the eyes to objects and words during processing of linguistic information related to these objects and words. Furthermore, do such fixations facilitate linguistic processing? I was especially interested in the role of the spatial location at which objects are positioned and how spatial location is represented within the cognitive system.

This thesis was divided into three main parts. In the first part, I investigated which role spatial location information plays in processing linguistic information in both language production and language comprehension. In order to investigate the role of the spatial location independently from the visual information conveyed by the objects, the objects were either completely removed from the visual field or changed their position. Former studies established that people look at empty regions previously occupied by objects, when processing referents of these objects (e.g. Altmann, 2004; Richardson & Spivey, 2000; Spivey & Geng, 2001). This effect was investigated more closely and utilised in order to examine whether fixating the relevant empty positions affects language production and processing.

In the second part, regressive eye movements during reading of garden path sentences were examined. Most central was the question whether readers are able to directly return to ambiguous regions of the garden path sentences in order to reanalyse the sentences. Are readers able to utilise the spatial location representation of words in the sentence to target critical words within the sentence?

The third part discusses the results and the implications for a model that integrated visual and linguistic information.

In the remainder of this chapter, I introduce some general facts about eye movements, language production, and comprehension.

1.1 The eye as the window to the mind

The eyes are the only organs of our body with which we are able to process visual information. The most important part in order so see is the retina. It consists of two types of receptors, the cones, and the rods, which serve very different roles. The cones, primarily situated at the centre of the retina, are processing detail and acuity of a visual scene. The rods are mainly located at the periphery of the retina. Their main task is to detect movements. Furthermore, the rods are able to distinguish light from dark objects. Between the centre and the periphery of the retina, both kinds of receptors are present. This distribution of cones and rods on the retina is responsible for how humans perceive visual input. For example, when looking at a scene, we are not processing every part of the scene in the same way. Rather, the most accurate information is perceived around the point of fixation. The area on the retina that processed the information from this location is called the fovea and extends to about two degrees of visual angle around the fixation point (Rayner & Pollatsek, 1989). The density of cones is the highest in this area. The area around the fovea is called the parafovea and extends to up to ten degrees of visual angle. Information is still processed in the parafovea but is not as accurate as in the fovea. The area beyond the parafovea is termed the peripheral area and only information from highly salient objects or movements is perceived. In this part of the retina mostly rods are present.

Since only the foveal part of the visual field is perceived in detail, people have to move their eyes around in order to get accurate information about multiple parts of a scene. This movement is not continuous and smooth but consists of single fixations that are interleaved by saccades. Within a fixation, the eyes remain relatively stable in one given location for about 250ms (Rayner & Pollatsek, 1989). This fixation duration is only a rough average, and the actual length of an eye movement depends on the task a person has to perform. In eye movement research, fixation time is one of the most important dependent variables in order to investigate cognitive processes. A saccade is a very fast eye movement (about 500°/s) and during a saccade no new visual information is perceived (Irwin, 2004; Matin, 1974).

An important point to mention is that the brain does not immediately process the information perceived by the eyes. There is a so-called *eye-to-brain lag* which is approximately 50ms long (Clark, Fan, & Hillyard, 1995; Foxe & Simpson, 2002; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000; VanRullen & Thorpe, 2001). However, since the eye-to-brain lag is very short, measuring eye movements is a powerful tool in order to investigate online cognitive processes (Just & Carpenter, 1980), and it can be assumed that a fixated region in the visual field is the region which is currently processed. This is called the *eye-mind assumption* (Just & Carpenter, 1978) and can be applied in visual tasks such as reading or visual search (Just & Carpenter, 1976).

After having summarised some of the principles of how the eyes work, I turn to eye movements in language production and comprehension. After a short introduction to how humans produce and comprehend language, eye movements during production and comprehension are laid out.

1.2 Language production

Language production is the process of generating spoken utterances. This process is not limited to only the articulation of words or sentences but includes the complete process from creating a concept about what a person wants to utter until the realisation of the phonetic code. Before describing the relationship between eye movements and language production, I first lay out a general model of language production introduced by Levelt (1989) and Levelt, Roelofs, & Meyer (1999).

1.2.1 How do speakers plan and produce utterances?

According to this model, language production is a staged and feed-forward process. Its main stages include a conceptual preparation, followed by lexical selection, morphological encoding, phonological encoding and syllabification, phonetic encoding, ultimately leading to an articulation stage. In the remainder of this section, I will shortly introduce each of these stages.

During the first stage, the *conceptual preparation stage,* a lexical concept is activated. This lexical concept is activated as a part of the communicative goal of the speaker. According to Levelt's model, the conceptual vocabulary does not consist of primitive conceptual features that are assembled into the lexical concept (Bierwisch & Schreuder, 1992; Goldman, 1975; Morton, 1969). The conceptual vocabulary rather consists of more complex entities that make the message, intended for production, more explicit (see e.g. Fodor, Garrettt, Walker, & Parkes, 1980). The lexical concept is forwarded to the *lexical selection* stage in which the lemma is retrieved from the mental lexicon. A lemma refers to the semantic-syntactic entity of a word (Kempen & Hoenkamp, 1987). The retrieval of a lemma makes a word's syntactic properties available. For example, retrieving the lemma *to give* makes the information available that *give* is a ditransitive verb that takes a direct and an indirect object. Furthermore, lemmas consist of diacritic parameters that have to be set at retrieval. These parameters include information about the number, person, tense, or mood of a lemma. The lemma is passed to the *morphological encoding* stage. This step marks the crossover from the conceptual towards the articulatory motor system, from the lemma towards the word form or lexeme. In this stage the morpheme is retrieved from the mental lexicon and then transferred to the next stage: the *phonological encoding* stage, in which the phonological word is retrieved. These two stages are often combined into the morphophonological encoding stage in which the syllabic and phonetic properties of a word are accessed. This information is passed towards the *phonetic encoding* stage in which the gestural scores are produced and finally articulated.

In addition to these stages, Levelt (1989) proposed a self-monitoring module. Speakers are able to monitor not just the articulated speech but also some internal representations. Levelt suggested that the conceptual preparation stage and the phonological word are subject to monitoring.

The process of language production is both *cascading* and *incremental*. This means that the processes always occur in a pre-set order such as outlined above. Furthermore, as soon as a stage has processed some information, it passes it immediately to the next stage before completely processing all information available.

Before turning to the eye movements during language production, I am going to describe the difference between word frequency and codability and which parts of the language production system they affect. The frequency of a word is defined by its occurrences in speech. For example function words like *the* occur very often and are of high frequency. Words like *lizard* occur less often and are thus of lower frequency. Word frequency affects the language production process such that it takes longer to name pictures depicting a low frequency object than it takes to name high frequency objects (Oldfield & Wingfield, 1965). This effect originates from slower access of the phonological form at the word form level when a word is of low frequency (Jescheniak & Levelt, 1994).

Codability is a measure of the number of possible names that are available to describe an object. The more likely it is that speakers name an object by using the same word, the more codable an object is. For example, a banana is always a banana and thus highly codable, whereas a suitcase could also be described as a trunk and is thus of lower codability. It has been shown that people name objects with a higher codability value faster than objects with a low codability value (Lachman, 1973; Lachman & Lachman, 1980). This effect is not due to a faster recognition time but due to a faster lemma selection of high codable objects (Johnson, 1992). This stands in contrast to word frequency which affects retrieval at the word form level (Jescheniak & Levelt, 1994).

1.2.2 Eye movements towards to-be named objects

To name an object, speakers first have to fixate this object in order to identify it. Although this statement sounds trivial, the time speakers actually spend on to-be named objects is worth investigating more closely, because it provides insights into how words are retrieved from the lexicon and how the language production system works in general.

People are able to rapidly recognise objects, or perceive the content of a scene (Bock, Irwin, Davidson, & Levelt, 2003; Potter, 1975). Potter (1975) showed that only 170ms of viewing time is necessary to categorise objects. Bock, et al. (2003) showed that people are able to tell the time after only 100ms of exposure to a clock. To recognise the gist of a scene (e.g. to say that a scene depicts a kitchen), only one fixation is necessary (Biederman, 1972; Biederman, Mezzanotte, & Rabinowitz, 1982). However, when naming objects or talking about the content of a scene, people fixate objects that correspond to their speech for much longer. These so-called *name-related gazes* are over 450ms (Meyer, et al., 1998) or even much longer (Griffin, 2001). In a study conducted by Griffin & Bock (2000), participants were instructed to describe scenes depicting a simple transitive event (e.g. *The mailman is chasing the dog)*. The authors showed that participants fixated objects within the scene (e.g. mailman, dog) in order of mentioning. The gaze duration of fixations at objects was over 800ms. Griffin & Bock also reported the so-called *eye-voice span*. This span denotes the time between the onset of the last gaze towards a to-be named object and the speech onset of the noun denoting this object. The eye-voice span in Griffin & Bock's experiment was approximately 900ms. In other words speakers start looking at objects about 1s before naming them. This is approximately the time people need to name isolated objects in a picture-naming task (Snodgrass & Yuditsky, 1996).

Name-related gazes vary according to multiple factors. Griffin (2001) investigated the duration of name-related gazes towards objects that exhibit different codability values and different word frequencies. In this study, participants were instructed to describe the spatial relation of three objects depicted on a computer screen. The goal of the study was to determine whether codability and frequency of objects affect the duration of speakers' gazes at these objects during naming within complex utterances (see Meyer, et al., 1998 for a similar study investigating naming of object pairs). Name-related gazes were longer when speakers named low codable words. Furthermore, gazes were also longer for low frequency objects. Since codability is associated with the retrieval of the lemma and word frequency with the retrieval of a word's phonological form, name-related gazes are associated with these two processes. The amount of time speakers spend on looking at objects thus, reflects the time it takes to retrieve the lemma and the phonological form of a word, and speakers keep looking at objects until the phonological form is retrieved (Meyer, et al., 1998). These results are supported by the finding that speakers do not fixate objects during articulation but move their eyes to the next to-be produced object at about 100-200ms before speech onset (Griffin, 2001; Meyer, et al., 1998).

In the previous paragraphs, I described name-related gazes (i.e. gazes that are directly associated with production of objects). However, in order to produce an object, speakers also have to identify it. For complex utterances, Griffin & Bock (2000) distinguished two stages in the language production process: the non-linguistic apprehension and the linguistic formulation phase. In the apprehension phase, speakers perceive the overall gist of a to-be described scene. This process is very rapid (300-400ms) and highly parallel. Only one eye fixation might be sufficient, and speakers do not need to fixate every aspect of the scene directly. In the following formulation phase, speakers look at single objects within the scene in a serial manner and according to the word currently produced. The apprehension phase corresponds to the conceptual preparation phase and the formulation phase to the following grammatical encoding (Bock, et al., 2003).

1.3 Language comprehension

After summarizing how people speak and how they move their eyes during language production, I now turn to the listener and language comprehension. I first give a short introduction to the process of language comprehension and models of sentence processing. I conclude this section by introducing the literature of eye movements during comprehension of spoken language and during reading.

1.3.1 How do listeners comprehend language?

At first glance, language comprehension is the reversed process of language production. In language production, a concept has to be transformed into an utterance, in language comprehension an utterance or a written sentence has to be transformed into a concept. However, there are a number of non-trivial differences. In the following paragraph I give a short overview of language comprehension.

In both spoken language understanding and reading, at first, single words have to be identified. In reading this process is relatively straightforward. In English, each letter is a single unit and words are usually divided by spaces. Thus, segmenting a sentence into its words is easy. However, in spoken language comprehension, a number of problems occur. First, speech has to be filtered from other signals like background noise and segmented into phonemes, which are the smallest units in speech. Second, the phonemes have to be assembled into single words. This procedure is not trivial. When speaking, people do not always pause between words such that the listener would be able to determine when a word ends and another begins. This is called the *segmentation problem*. A famous example is the phrase *I scream* which is acoustically not distinguishable from *ice cream.* There are a number of constraints and strategies that a listener can apply in order to overcome this problem. For example, the *possible word constraint*, which states that every part of speech should be attached to a word (Norris, McQueen, Cutler, & Butterfield, 1997). Another strategy that can be applied is the *metrical segmentation strategy*. A particular language has a characteristic way of structuring and segmenting sounds. For example, in English, strong syllables are usually word initial in content words, whereas weak syllables are not word initial (Cutler & Butterfield, 1992).

These strategies help the listeners to map the acoustic signal onto a representation of a word in the mental lexicon. It is assumed that this process starts immediately when the first few phonemes of a word are retrieved (e.g. Marslen-Wilson & Welsh, 1978). Thus, a first set of phonemes immediately activates a set of words. This set is reduced when more phonemes are retrieved and eventually only one word remains in the list. Accessing a word in the mental lexicon also accesses its morphological, syntactic and semantic information. This information is used to integrate the word into the larger sentential structure. A sentence is then integrated into a discourse model of the listener.

In the following section, I present an overview of different models of parsing. Furthermore, I introduce the notion of ambiguity which is central to Part II of this thesis.

1.3.2 Models of sentence processing

Parsing is the process of recovering the syntactic structure of a sentence. This process involves assigning the part of speech to each word in a sentence. The part of speech denotes to which lexical category (e.g. noun, verb, adjective) a specific word belongs. In a next step, these categories are combined into larger phrases such as a noun phrase (NP) or verb phrase (VP). Examples of phrase structure trees of a sentence can be seen in Figure 1.1.

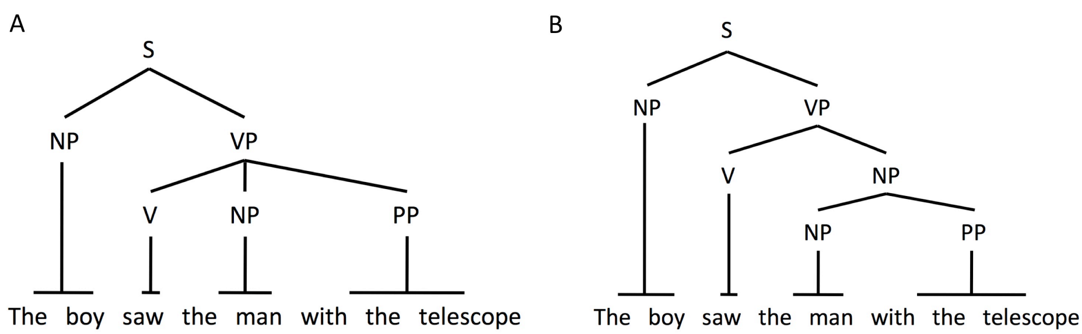

Figure 1.1: Two alternative phrase structure trees of the ambiguous sentence: *The boy saw the man with the telescope.*

This phrase structure is used to assign syntactic categories like subject or object, and the thematic roles to nouns and noun phrases. Thematic roles determine the meaning of a noun or noun phrase in terms of the action or state it takes within the sentence. For example, in the sentence in Figure 1.1 A, the noun phrase *the boy* denotes the agent and the noun phrase *the man* the theme of the sentence.

However, the structure in Figure 1.1 A is not the only interpretation of this sentence. According to the syntactic structure in Figure 1.1 A, the boy is using the telescope in order to see the man. Another interpretation would be that the boy sees a man who possesses a telescope (Figure 1.1 B). This example illustrates a sentence containing a so-called *global ambiguity*. In a globally ambiguous sentence, the ambiguity is not resolved and the sentence remains interpretable in more than one way.

A second type of ambiguity is the *local ambiguity.* Sentences including a local ambiguity can initially be interpreted in more than one way. However, the ambiguity is always resolved later in the sentence by the *disambiguating word*. A famous example is the following Sentence (1.1).

(1.1) *The horse raced past the barn fell.* (Bever, 1970)

In the initial analysis, the noun phrase *the barn* is directly attached as an object of the verb. However, in an alternative interpretation, the sentence contains a reduced relative clause (Sentence (1.2)).

(1.2) *The horse <u>that was</u> raced past the barn fell*.

The sentence is disambiguated when processing the word *fell*. This example demonstrates a special type of local ambiguity and such sentences are termed *garden path sentences*. In garden path sentences, the incorrect interpretation is initially the preferred one but when processing the disambiguating word the sentence has to be reanalysed. Garden path sentences are of high interest in parsing research because they allow an investigation of the human parser when it fails. In the next section, I am going to introduce some of the main approaches in human parsing that examined how syntactic ambiguities are processed and resolved.

Two main approaches have dominated the research on parsing: autonomous two-stage models and interactive one-stage models. In autonomous two-stage models, syntactic and semantic information are processed in separate steps and are not interacting, whereas in interactive one-stage models syntactic and semantic information are processed together.

1.3.2.1 Autonomous two-stage models

The garden path model (Frazier, 1987a) is an example of a two-stage model. In the first stage, only syntactic information is taken into account in order to parse a sentence. In case a sentence is ambiguous, only one interpretation is followed guided by two principles; the *minimal attachment* and the *late closure* principles. The minimal attachment principle states that a parse tree using fewer nodes, and thus syntactically less complex, is preferred towards a tree with more nodes. For example, in the globally ambiguous sentence in Figure 1.1, *The boy saw the man with the telescope,* the parse tree in which the prepositional phrase *with the telescope* is directly attached to the verb phrase (Figure 1.1 A) uses fewer notes than when it is attached to the noun phrase *the man* (Figure 1.1 B) and thus represents the preferred interpretation. To satisfy the late closure principle, new input has to be attached at the last processed phrase in the parse tree. Thus, in the previous example, the interpretation in which the propositional phrase is directly attached to the previous noun phrase is the preferred one. When there is a conflict between the two principles, the minimal attachment principle takes priority. If parsing of a sentence was not successful in this first stage, a second stage is initiated. In this stage semantic information like thematic role information is taken into account. Two-stage models predict that the two stages do not interact with each other.

There are a number of pieces of evidence favoring the two-stage garden path model (e.g., Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; Mitchell, 1987; Van Gompel & Pickering, 2001a). For example, Ferreira & Clifton (1986) showed that semantic information cannot prevent the parser from initially getting garden-pathed. Ferreira & Clifton presented participants with sentences exhibiting syntactic structure ambiguities (Examples (1.3) and (1.4)).

(1.3) *The <u>defendant</u> examined by the lawyer turned out to be unreliable.*

(1.4) *The <u>evidence</u> examined by the lawyer turned out to be unreliable.*

According to minimal attachment, the verb *examined* should initially be interpreted as being the main verb instead of a verb in a reduced relative clause. However, in some of the sentences (Example (1.4)) the animacy of the subject disambiguated the sentence early. Since *evidence* is an inanimate noun, it cannot serve as the agent of the sentence and the verb must be therefore part of a reduced relative clause. However, reading times suggested that readers were still garden-pathed and that the animacy information was not included into the syntactic analysis.

1.3.2.2 Interactive one-stage models

An example for the interactive one-stage model is the *constraint-based* model (e.g. Boland, Tanenhaus, & Garney, 1990; Tanenhaus, Carlson, & Trueswell, 1989; but see). According to this model, multiple interpretations of a sentence are built in parallel. However, based on syntactic, semantic and discourse constraints one interpretation is favoured. The parser is led into a garden path when the correct analysis is not preferred by these constraints.

There are many studies showing that the parser initially does not solely rely on syntactic information (e.g. Altmann & Steedman, 1988; Milne, 1982; Tanenhaus, et al., 1995; Trueswell & Tanenhaus, 1994; but see e.g. Mitchell, Corley, & Garnham, 1992). For example Altmann & Steedman (1988) presented participants with ambiguous sentences such as Sentence (1.5). In this sentence, the prepositional phrase *with the dynamite* can either modify the noun phrase *the safe* or the verb phrase *blew open the safe*.

(1.5) *The burglar blew open the safe with the dynamite and made off with the loot.*

However, the authors provided referential context that disambiguated the sentences early. For example, the verb phrase interpretation is preferred within the context: *A burglar broke into the bank carrying some dynamite*. Response time data showed a significant influence of the referential context suggesting that this information is used to disambiguate ambiguous sentences. Trueswell & Tanenhaus (1994) argued that the reason why Ferreira & Clifton (1986) did not find any effects was because the semantic bias in the presented sentences was too weak to cause an effect.

1.3.2.3 Other models

A more recent parsing model is the unrestricted-race model (Traxler, Pickering, & Clifton, 1998; Van Gompel & Pickering, 2001b; Van Gompel, Pickering, & Traxler, 2000). According to this model both syntactic and semantic information are used. However, alternative structures are constructed in parallel and "race" against each other, and the interpretation constructed fastest wins the race. Only one interpretation remains and if this is an incorrect one, the parser has to reanalyse a sentence.

The good enough theory (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, 2003) follows a different approach. It assumes that the parser does not always engage in a complete analysis. People use simple heuristics in addition to detailed analyses. As a result, if a sentence was not correctly parsed, people might not engage in a reanalysis and retain an incomplete and incorrect interpretation.

The models of parsing presented in this section are far from being complete. They only provide an overview of some of the main directions research on parsing took in the past years. In the second part of this thesis, I will return to the topic of parsing but concentrate more on how readers recover from miss-parsed sentences.

1.3.3 Eye movements and comprehension

This section summarises the literature about eye movements and spoken language comprehension. I only introduce the most important literature. For a more detailed description refer to the Chapter 2.

In a very influential study, Cooper (1974) found that people attend to objects in their visual field that are closely related to words currently heard in unfolding speech. Cooper (1974) showed participants a set of nine objects in a three-by-three grid while they listened to a short story. People were more likely to look at a lion than to e.g. a dog when hearing the word '*lion*' and at a snake or a zebra when hearing the word '*Africa*'. This method that exploits the tendency to fixate objects at approximately the same time as words related to this object occur in speech was termed the *visual world paradigm*, and was utilised in numerous psycholinguistic studies to investigate linguistic processing.

Tanenhaus, et al. (1995) investigated eye movements towards multiple objects while listening to complex sentences. Participants were presented with instructions like Example (1.6) and with a set of real objects that included objects from the sentence (i.e. an apple, an apple on a towel and an empty box).

(1.6) *Put the apple on the towel in the box.*

The task was to follow the instruction. Of main interest were eye movements during comprehension of the instructions. The authors found that people fixated the objects they heard in the instruction at approximately 250ms after the offset of the word that specified this object. These results were taken as evidence for an incremental and immediate integration of the visual world with spoken language. They furthermore showed that the visual world is used in order to disambiguate language. In the same set up, participants saw sets of objects that included either an apple on a towel, an empty towel and an empty box (set-up A) or an apple on a towel, an additional apple on a napkin, an empty towel and an empty box (set-up B). Participants were presented with instructions like in Example (1.6). The instructions exhibited a local ambiguity. In Example (1.6), before hearing the word *in*, the prepositional phrase *on the towel* could either specify the location in which the apple has to be put or it could modify the noun phrase *the apple*. Tanenhaus, et al. (1995) found that participants fixated the empty towel only in set-up A which included only one apple. In set-up B, the phrase p*ut the apple* is underspecified because the listener does not know which apple needs to be used; either the apple on the towel, or the apple on the napkin. Thus, participants expect a modifying phrase and do not get garden-pathed. The results of this study provide evidence for an immediate integration of visual information into syntactic processing and are proof against autonomous two-stage models.

The integration of visual information is even anticipatory. It has been shown that people attend to related objects before they occurred in speech (Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003). Altmann & Kamide (1999) showed participants a scene including a boy and some objects like a cake, a ball, and a toy train. When people heard a sentence as in Example (1.7), they already looked at a cake while hearing the word '*eat*'.

(1.7) *The boy will eat the cake*.

The cake was the only edible object in the scene. Participants utilised this information and were able to integrate it very early into their linguistic processing mechanisms.

Participants also attended to objects that share visual features with objects that are a direct referent of an unfolding word (Huettig & Altmann, 2007; Myung, Blumstein, & Sedivy, 2006). In an experiment by Huettig & Altmann (2007), participants heard words like '*snake*' while looking at a set of objects. The authors showed that people's attention moved to a picture of a cable as its shape properties matched the activated shape properties of a snake.

In summary, people attend to objects with identical or similar visual features whilst hearing or anticipating words related to these objects. In Chapter 2, I will discuss this issue more closely. In particular, the phenomenon that people not only attend to objects but also to positions at which objects were located when hearing or anticipating words related to these objects is discussed.

1.3.4 Eye movements in reading

In this section, I am going to introduce how readers move their eyes in order to read and process written text. It is believed that reading and spoken language recognition develop very closely. For example, comprehension skills in both systems show a high correlation (Palmer, MacLeod, Hunt, & Davidson, 1985). However, there are differences, and I first summarise some of the differences between reading and spoken language understanding.

First of all, written text is available for much longer for readers in comparison to spoken language for listeners. As a result, readers can linger as long as necessary on single words. In reading it is also possible to regress to earlier parts of a sentence. In comparison, in order to recover spoken language, the listener would need to rely on memory, or ask the speaker to repeat an utterance. Furthermore, in written text, no prosodic information is available. In spoken language, prosody is used for syntactic parsing (Kjelgaard & Speer, 1999; Schafer, Carlson, Clifton, & Frazier, 2000), but it cannot be utilised during reading. These differences between reading and spoken language comprehension are also reflected in how sentences are parsed (Kennedy, Murray, Jennings, & Reid, 1989). In the following sections, I introduce how people move their eyes during the reading process. Most information is taken from Rayner & Pollatsek (1989).

As I explained in Section 1.1, the eyes do not move in a continuous fashion. Eye movements consist of single fixations that are connected by fast saccades. In reading, the average length of a fixation is approximately 200-250ms. A saccade is usually 7-9 characters long, and takes between 20 and 35ms. We are able to process about 200-300 words per minute depending on the task, the kind of text and age of the reader (Carver, 1990; Just & Carpenter, 1980; Potter, Kroll, & Harris, 1980; Rayner & Pollatsek, 1989).

During reading, people perceive information within the so-called *perceptual reading span* (McConkie & Rayner, 1976; Rayner, Well, & Pollatsek, 1980). The perceptual reading span defines the area in which effective visual processing during reading is possible. When reading English sentences, this area is asymmetric to the right, spanning 3-4 characters to the left and 12-15 characters to the right of a given fixation (McConkie & Rayner, 1976; Rayner & Pollatsek, 1989; Underwood & McConkie, 1985). The reason for this asymmetry is that in forward reading, before the eyes move to a new location, attention moves first (Henderson, Pollatsek, & Rayner, 1989). Morrison (1980) compared this process with the analogy of a rubber band. Attention moves first and the eyes are pulled behind like being attached to it with a rubber band. Moving attention extends the region of visual perception. The text is pre-processed and the information is used in order to plan a saccade to the next location (Rayner, 1975). In languages like Hebrew in which the normal reading direction is from right-to-left, the perceptual span accordingly extends to the left (Pollatsek, Bolozky, Well, & Rayner, 1981). Likewise, when readers of English read inverted English text from right-to-left, the perceptual span also extends to the left (Inhoff, Pollatsek, Posner, & Rayner, 1989).

Only very little information from outside of the perceptual span is acquired (Rayner & Bertera, 1979; Rayner, McConkie, & Ehrlich, 1978). Most of the studies, investigating the perceptual reading span, used the so-called *moving window technique* (McConkie & Rayner, 1975). While a participant reads through a sentence, letters to the left and right of the current fixation are replaced by Xs or different letters. By varying the number of unchanged letters around the point of fixation, it is possible to determine how many letters to the right and left are processed during a single fixation.

An important proportion of eye movements in reading are regressions. Regressions are saccades that move in the opposite direction of the generic reading direction in a specific language. In English, regressions are movements going from right-to-left. In normal reading of English text, readers initiate regressions in 10-15% of all eye movements (Rayner & Pollatsek, 1989; Vitu & McConkie, 2000). Readers regress for a number of reasons, such as in response to low-level visuomotor or word identification processes or due to higher-level syntactic and semantic processes (Vitu, 2005). Regressions due to visuomotor and word identification processes are usually very short and consist mostly of inter-word regressions or regressions that land on the word immediately to the left of the launch site. Regressions due to higher-level processes are longer, and the length of regressive saccades can cover multiple words (Frazier & Rayner, 1982; Meseguer, Carreiras, & Clifton, 2002). In the second part of this thesis, I investigate regressive eye movements while reading garden path sentences. Regressive eye movements are introduced in more detail in Part II.

1.4 Overview of the thesis

This thesis is concerned with the relationship between visual and linguistic processing. In more detail, I investigated the question of how people attend to objects in their visual environment, when processing linguistic information related to these objects. The role of the spatial location of objects received special attention, and I investigated the role of spatial location in the presence or absence of objects.

The thesis is divided into two main parts. In the first part, the interaction between visual and linguistic information was examined during language production and during processing of information conveyed by spoken language, which, in the remainder of the thesis, is also referred to as language comprehension. Within the first part, in Chapter 2, I introduced the literature and models concerning the interaction of visual and linguistic processing. The main focus is on the effect of *looking at nothing*, which shows that people fixate empty regions that were previously occupied by an object when processing information related to this object.

Chapter 3 explored the interaction between vision and language in language production. Two experiments were conducted. Experiment 1 investigated the role of spatial location by manipulating the position of to-be named objects. In Experiment 2, the objects were withdrawn during the language production process and I investigated the looking at nothing effect. Furthermore, influences of eye movements towards empty regions on language production were examined.

In Chapter 4, similar questions as in Chapter 3 were investigated during language comprehension. In three Experiments (Experiments 3-5), the role of spatial location was explored by moving objects during retrieval of information related to these objects (Experiment 3), and by presenting an empty screen (Experiment 4 & 5).

In the second part of the thesis, I examined regressive eye movements during reading of garden path sentences (Chapter 5). In Experiments 6 & 7, I investigated the utilisation of spatial location information of single words within sentences. How do readers regress to earlier parts of a sentence when reanalysing ambiguous sentential structures. I also examined the perceptual reading span before a regression was launched. I hypothesised that attention moves to the left before a regression is launched. The information perceived from this shift of attention might be utilised in order to guide regressive eye movements.

In Chapter 6, I summarised the results from both parts of the thesis and discussed the implications for a model that integrates both language and visual processing.

Part I: The integration of visual and linguistic information

2 Chapter 2: Introduction

Speaking and listening are not isolated tasks, decoupled from the visual environment. In the previous chapter, I presented studies showing that speakers gaze at objects they are about to name for much longer than is required for identification. This prolonged exposure to visual information supports the retrieval of a word's lemma and lexeme. Furthermore, listeners fixate objects in their visual environment mentioned in speech and incorporate this visual information in processing the syntactic structure of linguistic input.

Previous theories suggested that language and visual processing are administered by modules that are independent from other non-linguistic information (Fodor, 1983; Jackendoff, 2002). However, results of recent studies rather suggest that language processing is closely interconnected with the representations of visual information (e.g. Cooper, 1974; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Tanenhaus, et al., 1995). A theory which is often taken to support this proposal is the perceptual symbol theory (Barsalou, 1999). Within this theory, it is assumed that linguistic information is stored as perceptual representations. This represents an alternative to an amodal propositional structure theory (Kintsch & Van Dijk, 1978), which assumes that linguistic information is stored as propositional representations such as [EAT[BOY, CAKE]]. Within the perceptual symbol theory, the linguistic input is stored as perceptual symbols which closely resemble visual information. During language comprehension, perceptual symbols are activated. As a result, the visual information becomes part of the mental representation of linguistic input (Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002).

Zwaan, et al. (2002) illustrated this process in an experiment in which participants read sentences like (2.1) and (2.2).

(2.1) *The ranger saw the eagle in the sky*

(2.2) *The ranger saw the eagle in its nest*

Subsequently, participants were shown pictures and they had to decide whether the depicted objects occurred in the sentence. The authors found that reaction time was faster when the objects fit the context of the sentence. For example, after reading Sentence (2.1), people were faster identifying an object depicting an eagle with outstretched wings in comparison to one with tucked in wings. The eagle in the sky needs to have outstretched wings in order to fly. This feature of the eagle is part of the representation of the eagle in Sentence (2.1), and thus people are faster in the judging task.

This study showed that people activate perceptual representations of objects during linguistic processing. It also has been shown that people activate a representation of movement during language comprehension (Matlock, 2004; Richardson & Matlock, 2007). When participants are presented with scenes depicting a road in a desert, they tend to fixate the road more often when hearing sentences like (2.3) in comparison to sentences like (2.4).

(2.3) *The road runs through the desert.*

(2.4) *The road is in the desert.*

Sentence (2.3) describes the scene by using fictive motion. Fictive motion seems to be represented as a real motion that affects eye movements. Interestingly, when introducing a context in which the road passes through difficult terrain (e.g. *The desert is hilly*), people spend more time on the road in comparison to when it passes through easier terrain.

These studies reveal interesting aspects of how linguistic information is represented.

> … the representation of meaning from linguistic input is a dynamic process involving malleable perceptual representations rather than the mechanical combination of discrete components of meaning (Zwaan, et al., 2002, p. 170).

These studies present an explanation for the findings that people gaze at objects during language production and comprehension. If the representation of linguistic information includes a visual component, looking at a relevant object during linguistic processing can support language production and comprehension. In the remainder of this chapter, this issue is discussed in more detail. In the next section, I illustrate the looking at nothing phenomenon that extends previous findings such that people even fixate the location in which objects have been during language production and comprehension. This phenomenon suggests that the spatial location of objects also interacts with linguistic processes related to these objects.

2.1 Looking at Nothing

In the current section, I outline experiments investigating the looking at nothing phenomenon (Altmann, 2004; Hoover & Richardson, 2008; Johansson, Holsanova, & Holmqvist, 2006; Knoeferle, Habets, Crocker, & Muente, 2008; Richardson & Spivey, 2000; Spivey & Geng, 2001). In Section 1.3.3, I introduced the phenomenon that people fixate an object when hearing a referent of this object in speech. However, if the relevant object has been removed just before an utterance including a reference of this object is processed, people also tend to look at the position that the object previously occupied.

In two experiments, Spivey & Geng (2001) investigated how people move their eyes when looking at an empty screen. They tested eye movements during imagining a scene (Experiment 1) and during remembering objects from a previously presented scene (Experiment 2). In Experiment 1, participants heard stories that were biased for a specific direction. For example, an "upward" story described an apartment building by naming the inhabitants starting at a lower floor and working its way up to the top floor. In a "leftward" story railway carriages of a train were described starting at the left side of the train. Spivey & Geng found that while looking at an empty screen, saccade movements corresponding to the directional bias of a story occurred more often than movements in other directions. In Experiment 2, participants first had to fixate four simple objects (e.g. triangle, square). Following this first phase, the screen went blank and people were asked about the colour of one specific object. During retrieval of the colour, participants were more likely (24% of the trials) to fixate the blank region in which the relevant objects used to be in comparison to looking at the position of irrelevant objects. However, in the majority of the remaining trials, the eyes remained at the centre of the screen.

Spivey & Geng's study investigated looking at nothing during imagining a scene and after memorizing a set of simple objects. Altmann (2004) tested eye movements with slightly more complex sceneries. Similar to the material in Altmann & Kamide (1999; see Section 1.3.3 for a description of this study), participants saw a scene depicting two people and two items. Furthermore, they were presented with an auditory stimulus describing the scene (e.g. *The man will eat the cake*). However, before speech onset of the stimulus, the scene disappeared and was replaced by a blank screen. Altmann found that people fixated the location in which an object had been during the auditory presentation of the referent of the object. Furthermore, as in Altmann & Kamide (1999), participants even anticipated an upcoming word. For example, they attended to the position in which the cake had been when hearing the word '*eat*' because cake was the only edible object in the previously seen scene.

People also look at nothing during retrieval of contextual information associated with a specific location (Richardson & Spivey, 2000). In a first experiment of a set of five, Richardson & Spivey presented participants with talking heads that appeared consecutively in the four ports of a two-by-two grid. The heads presented four different facts (e.g. Sentence (2.5)) in each of the four ports.

(2.5) *Shakespeare's first plays were historical dramas; his last was the Tempest*.

Afterwards, participants heard a test fact which referred to one of the four introduction facts, while looking at an empty screen (the two-by-two grid was still visible). The participants had to evaluate the test fact to be either true or false in relation to the referring introduction fact. Richardson & Spivey (2000) found that while evaluating the test fact, participants attended to the corresponding port in which the referring fact was introduced more often than to other ports. In Experiments 3 and 4 participants were asked to keep their eyes at the middle of the screen during the presentation of the facts. In addition, in Experiment 4, it was ensured that although the fixation remained at the centre of the screen, attention moved to the corresponding ports in the two-by-two grid. In these experiments, the effect of looking at the corresponding ports while evaluating the test facts was not found. Richardson & Spivey (2000) suggested that attention to a port alone is not sufficient but that the effect depends on an initial fixation at the position in which the introduction fact was presented. In all five experiments, no facilitation was found when participants fixated the relevant empty ports in comparison to when they did not fixate it. Similar results were found by Hoover & Richardson (2008).

To summarise, people fixate empty locations when processing information associated with an object that formerly occupied this location. Fixations are triggered during imagining a novel scene, during remembering single objects and even during retrieval of contextual information related to an object. In the next section, I introduce models that attempt to explain the looking at nothing effect.

2.2 Models of integrating visual and linguistic information

2.2.1 An externalism approach

A very appealing explanation of the looking at nothing effect is the external world account. In contrast with an internalism approach, externalism does not assume that the mind can be described by the state of the brain alone. It rather postulates that the external environment is an integral part of mental states (see Spivey, Richardson, & Fitneva (2004) for a review). Thus, not all information of for instance an object has to be stored internally but can be outsourced in the external world and “the outside world serves as its own, external representation” (O'Regan & Nöe, 2001). In order to retrieve this information only so-called deictic pointers are stored internally (Ballard, Hayhoe, Pook, & Rao, 1997). These pointers act as an address to the external information, and by moving the eyes to this address external information can be easily retrieved. An advantage for the cognitive system is that working memory resources can be conserved because only a fraction of the actual information has to be stored.

According to this model, why do people look at empty locations when retrieving information associated with this location? During initial encoding of visual information from objects (Altmann, 2004; Spivey & Geng, 2001), or auditory facts (Richardson & Spivey, 2000), a deictic pointer is created that encodes the location of this information. During retrieval, this pointer becomes activated and the eyes move to the location of the object even if the object is not present anymore.

The change detection or change blindness phenomenon was taken as evidence for the external world account (Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1998). Change blindness is a phenomenon in which changes in a visual scene are not noticed even if they happen in a region of central interest (Rensink, et al., 1997). For example, Simons & Levin (1998) designed an experiment in which a confederate asked pedestrians for directions on a map. During the description, two further confederates, who carried a door, walked between the pedestrian and the first confederate, and thus, blocking the pedestrian's view of the first confederate. However, during that time, one of the confederates, carrying the door, swapped places with the first confederate and continued the conversation with the pedestrian. Intriguingly, the participant noticed the change in only about 50 percent of the trials. Early studies assumed that the change blindness phenomenon is due to the inability of the memory system to retain an internal representation of a scene. If a change happens outside a region of attention, it remains unnoticed (McConkie & Zola, 1979; Phillips, 1974; Rensink, et al., 1997). The pedestrian did not notice the change of the confederates because he had no internal memory of the first confederate.

However, the externalist approach has been undermined by a number of recent studies (Castelhano & Henderson, 2005; Henderson, 2008; Henderson & Castelhano, 2005; Hollingworth, 2008; Hollingworth & Henderson, 2002). Hollingworth & Henderson (2002) discovered that people are able to retain an internal representation of their visual environment. Whether people succumb to change blindness rather depends on whether a changing object has been initially fixated. Hollingworth & Henderson (2000) conducted an eye tracking experiment in which objects within a scene were changed after they had been fixated, but at a time when the focus of attention was not in the region of these objects. Participants were able to detect these changes.

People are even able to retain information about objects that were not intentionally memorized (Castelhano & Henderson, 2005). Castelhano & Henderson (2005) presented scenes in which participants had to perform visual search tasks. Afterwards, participants were asked to remember objects and the orientation of objects that were not the target in the search task. Participants were initially not aware that a memory task was part of the experiment. However, performance of the memory task was above chance level which indicates that people had built an internal representation of such objects.

The results from these studies indicate that information about objects is stored internally and besides spatial pointers towards objects, additional information must be stored. In the next section, I present further approaches that are able to explain the looking at nothing phenomenon.

2.2.2 An internalism approach

In this section, I introduce an alternative theory. In contrast to the externalist approach, this model includes a rich internal representation of visual information which is closely integrated with linguistic information. The model is closely related to the concept of objects files (Kahneman, Treisman, & Gibbs, 1992) and before discussing the model I briefly introduce object files.

2.2.2.1 Objects files

The concept of object files explains how the visual system is able to retain a stable representation of objects in the visual environment. Real world objects constantly change their spatial position, their relative position to the viewer and sometimes even briefly disappear behind other objects. Each of these changes produces a different projection of the object on the retina.

However, the viewer is effortlessly able to combine these projections into one united object representation. Object files were introduced as a “temporary ‘episodic’ representation of real world objects” (Kahneman, et al., 1992, p. 176). This representation is distinct from object representations in long-term memory in which information about the identity and classifications of objects are stored. Object files comprise a set of features of an object like colour or form, and can be accessed by their spatiotemporal correspondence. If an object is perceived, features of this object are encoded into an object file. Based on the spatiotemporal information of a stored object file, perceiving a new object or an object that changed its orientation causes the retrieval of a corresponding object file. If the spatiotemporal information matches, changed features are updated. Spatiotemporal information of an object file is updated when the object gradually changes its position. As an example for this process, Kahneman, et al. (1992) cited the famous American movie line “It’s a bird; it’s a plane, it’s Superman!”. While looking at this flying object, the spatiotemporal information and features of the object file associated with this object is constantly updated while Superman draws nearer. Importantly, activating an object file by its spatiotemporal information activates all other features of this object file. Kahneman, et al. (1992) presented evidence for this model in a number of experiments. In one experiment, participants were presented with a ‘preview screen’ consisting of a number of squares. Two of these squares contained letters, and the array was visible for 250ms (Experiment 1) followed by a screen in which all squares returned to being blank. This screen was visible for 300ms. In the following target screen, one of the letters appeared either in the same square (Same object condition), in a different square (Different object condition) or a new letter appeared in one of the two squares. The task for the participants in this experiment was to name the letter in the target screen as fast as possible. Kahneman, et al. (1992) found that people were faster when the target screen letter appeared in the same square in comparison to a different square or when a new letter appeared. Thus, the authors showed that activating the spatiotemporal information of an object file is essential in order to retrieve other features of the object file.

2.2.2.2 A model of integrated linguistic and visual information in memory

In order to model an integration of language and visual processing, Altmann & Kamide (2007) suggested a model based on four assumptions which are related to the idea of object files. First, objects and words are encoded into internal representations that consist of multiple features including information about location, colour, form, function, association, etc. of objects or words. The second assumption states that in case multiple representations with identical features are activated, the resulting coactivation of the same features leads to a boost of activation of the entire representation. Third, a change in featural activation is reflected by "changes in the attentional state of the cognitive system" (Altmann & Kamide, 2007, p. 512). As a consequence, the fourth assumption states that people are more likely to fixate objects with a higher featural activation.

Why do people fixate objects when hearing a referent of this object? It is important to mention that in experiments investigating the integration of language and vision, objects or the scene are usually shown before speech onset. Thus, when first looking at a picture of for example a snake, features like location, colour, form, etc. of the snake are encoded. The same features are encoded when people hear the word '*snake*' shortly afterwards, which causes a coactivation of the same features. The coactivated features boost the activation of the whole representation of the snake including the location which causes the eyes to revisit the picture of a snake. Likewise, when looking at a picture of a cable the form feature (e.g. wiggly) is being encoded. When hearing '*snake*' the same form feature (wiggly) is coactivated which causes the eyes to fixate the cable. Thus, not all features need to be coactivated but a subset is sufficient in order to induce a saccade to an object. Finally, looking at regions in which an object has been when hearing a referent can be explained by the same mechanism. When hearing a referent of an object which was visually presented just before the auditory presentation, the features of the objects are activated which results in a boost of activation of the whole representation, including the location. Thus attention shifts to the location of this object even if it is no longer visible.

The theory by Ferreira, Apel, & Henderson (2008) follows a very similar approach. However, the authors introduced a model that explicitly predicts that attending to the location of an object that has previously been encoded facilitates retrieval of linguistic information related to this object, even if the object is no longer visible. Similar to Altmann & Kamide (2007), Ferreira, et al. (2008) adopted the assumption that representations of objects consist of various features. These features are interconnected and activation of one feature activates other features related to the object (see Figure 2.1).

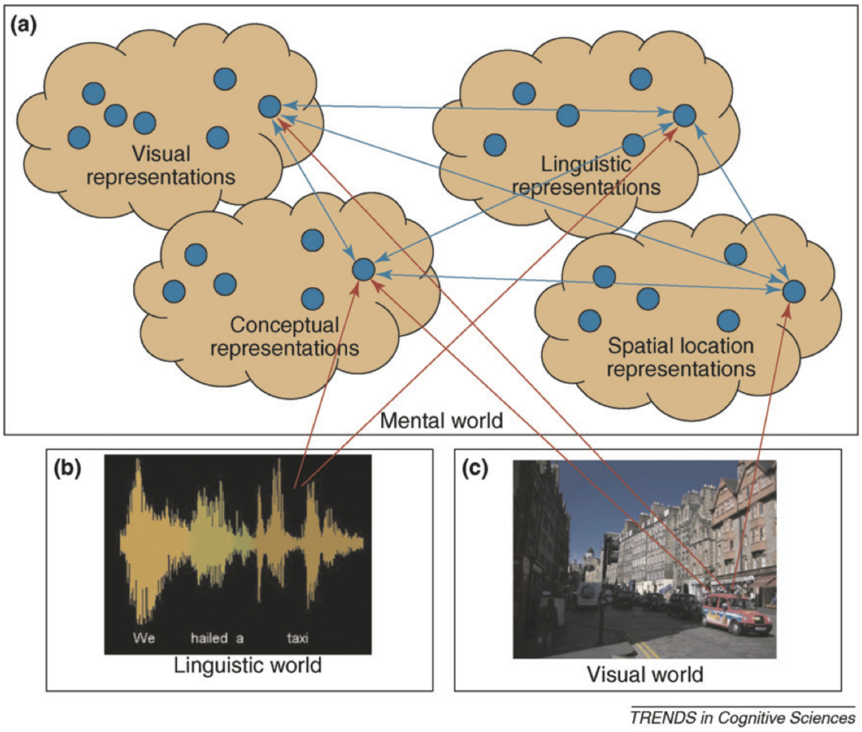

Figure 2.1: A model of integrated linguistic and visual information in memory (taken from Ferreira, et al. (2008)). Figure (a) depicts the mental world in which the visual, linguistic, conceptual and spatial location representations are closely linked. Information from the linguistic world (b) informs the linguistic and conceptual representations and information from the visual world (c) informs the visual, conceptual and spatial location representations.

The features are combined into different types of representations that the cognitive system can form, such as visual, linguistic, conceptual and spatial location representations[1]. When people are presented with a scene of a man and a cake, they form a representation that includes visual, linguistic, conceptual and spatial location information of the scene including the man and the cake. When they hear the sentence 'the man will eat the cake', the linguistic representation of e.g. the word 'cake' will also activate all other representations including the spatial location representation. The activated spatial location representation causes the eyes to move to the location of the cake in the scene even if the cake is not visible anymore. The theory also predicts that by moving the eyes to this location, the repeatedly activated spatial location representation also activates the other representations such as linguistic or conceptual representations, which in turn facilitate processing of linguistic or conceptual information.

Evidence that object memory performance is related to the location of this object also comes from the visual cognition literature (Hollingworth, 2006; Hollingworth & Henderson, 2002; Laeng & Teodorescu, 2002). Hollingworth & Henderson (2002) investigated the change blindness phenomenon. As mentioned above, people are sometimes blind to sudden changes in their visual field and the authors argued that this effect is due to the fact that the changed object was not initially encoded. However, Hollingworth & Henderson (2002) found that people almost always refixated the changed object when detecting the change; an issue also examined in this thesis. Hollingworth & Henderson argued that:

> "… refixation appears to play an important role in the retrieval of a stored object" (Hollingworth & Henderson, 2002, p. 131).

Hollingworth (2006) tested visual memory for objects within a scene when the objects changed their location in comparison to when they remained

---

[1] The 'spatial location representation' term in this thesis is equivalent to the spatiotemporal information in the Kahneman, et al. (1992) paper and in the remainder of this paper we will use the term from Ferreira, et al. (2008).

at the same position. He found that memory for the orientation and specific details of an object was better when the object was encoded in its original position in comparison to a different location. Similar results were found in real world scenes (Hollingworth, 2003, 2005).

However, a number of studies also reported that returning to objects does not facilitate the performance of object related tasks (Glenberg, Schroeder, & Robertson, 1998; Richardson & Spivey, 2000). Richardson & Spivey (2000) showed that attending to an empty space that an object previously occupied does not facilitate retrieval of information related to that object. The experiment by Richardson & Spivey (2000) was described in Section 2.1. As mentioned before, they also compared the accuracy of the evaluation when participants looked at the corresponding empty ports in comparison to when they did not look. The authors did not find any significant differences.

However, a problem with Richardson & Spivey's design was that items in which participants fixated the corresponding empty port and items in which they did not fixate it had to be selected post-hoc. Thus, items might not have been equally distributed between difficult and easy items. It is possible that ports previously occupied by difficult items were fixated more often than ports previously occupied by easy items. According to the theory by Ferreira, et al. (2008), fixating such empty ports facilitates retrieval of information. Due to an increased number of fixations, difficult items could thus be more facilitated than easier items. This trade-off effect between fixation and difficulty could therefore have been responsible for the reported null effect (see however Richardson, Altmann, Spivey, & Hoover, 2009).

In the following two chapters, the models by Altmann & Kamide (2007) and Ferreira, et al. (2008) are tested in language production and comprehension. The central questions of the experiments were whether people fixate empty positions, previously occupied by an object, when producing or comprehending these objects. Furthermore, do people benefit from moving their eyes to these empty regions?

Before continuing to the first experimental chapter, I will shortly discuss the power of the planned experiments in predicting eye movement and cognitive behaviours. A main aim of the experiments was to determine whether eye movements towards a specific location influence the performance of tasks related to objects that are or were located at this location. However, the aim could be formulated in a reversed order. Does linguistic performance related to objects in a specific location influences eye movements towards this location.

The experiments of the present study are designed such that eye movements are not directly manipulated. The reason for this design was that I did not want to introduce additional visual stimuli to trigger eye movements towards a specific location. A new visual stimulus might disrupt the influence of the experimental stimuli.

The advantage of this design is that I could directly investigate the relationship between eye movements and performance in object related linguistic tasks. However, the disadvantage was that it is difficult to argue about the causal relation of eye movements and performance. Does an object that is difficult to process cause the eyes to move to this object, or do people intentionally fixate difficult items in order to support processing?

The current thesis does not answer this question. However, I present evidence that fixating an object facilitates processing of linguistic information related to this object. Thus, there is a clear relationship between eye movements and performance. Further research is necessary to investigate the direction of causation of this relationship.

3 Chapter 3: The integration of visual and linguistic information in language production

3.1 Introduction

Section 1.2.2 provided a general overview of eye movements during language production. Speakers gaze at objects during preparation of the object's name for over 500ms (Griffin, 2004b; Meyer, et al., 1998). However, for mere identification of an object only 150ms viewing time is necessary (Potter, 1975).

In the current chapter, I examined more closely why speakers fixate objects for much longer than necessary for mere identification. I hypothesised that these name-related gazes are a consequence of an integrated model of vision and language. I was particularly interested in the role of the spatial location of to-be named objects. Similar to previous studies (Bock, et al., 2003; Griffin & Oppenheimer, 2006), I hypothesised that fixating an object during language production facilitates the language production process.

Bock, et al. (2003) proposed that looking at objects during language production supports the retrieval of information from memory. When encoding a scene in the apprehension phase, its content must first be memorised in order to talk about it later in the formulation phase. When retrieving this information the eyes attend to the corresponding referents in the scene to ease retrieval and access of the corresponding lemmas and lexemes.

Griffin & Oppenheimer (2008) proposed a similar theory. However, they further investigated what kind of information is retrieved during name-related gazes. Griffin & Oppenheimer (2008) introduced two contrasting hypotheses: the *content hypothesis* and the *referential hypothesis*. According to the content hypothesis, speakers fixate objects in order to continue processing the visual information of the objects. The referential hypothesis proposes that only the spatial location information of objects drive the eyes to the position of the objects.

> "Rather than mediated by the visual-conceptual features of the referent object, gaze is mediated by the spatial index of the referent. In other words, the intention to talk about an object orients a speaker's visual attention to the object's spatial location while processes related to the object are carried out." (Griffin & Oppenheimer, 2006, p. 944; see also Griffin, 2004b).

This theory is related to the objects file theory (Kahneman, et al., 1992; see also Section 2.2.2.1). During retrieval of an object from memory, the spatial location feature of this object is activated which leads the gaze to fixate the location referred to by the spatial index.

To investigate these two contrasting hypotheses, Griffin & Oppenheimer (2006) conducted an experiment in which speakers had to name objects with either their proper names (e.g. *dog* for a picture of a dog) or with a novel, unrelated word (e.g. *blick* for a picture of a dog). The content hypothesis predicts that when naming an object with an unrelated word, the visual features of the object would intervene with the retrieval of the novel word. Thus, speakers would be less likely to gaze at the object, or even look at neutral content. The referential hypothesis predicts that speakers fixate the object because it is still this object that has to be named. The spatial index refers towards this object which causes the eyes to fixate the object. The visual features of the object however, do not interfere with the language production process.

The authors found no differences in gaze duration at objects or in speech onset times between when the objects had to be named with their proper or with unrelated names. Griffin & Oppenheimer (2006) interpreted these results as evidence for the referential hypothesis. However, they assumed that both the content and the referential hypotheses are not mutually exclusive. Visual information of objects might still be retrieved and utilised in order to support the language production process.

In the current study the general idea of this theory is adopted. However, I propose a more general connection between memory and the language production system. As mentioned in Section 2.2.2.2, Altmann & Kamide (2007) proposed a cognitive system in which visual, linguistic, conceptual and spatial location representations are tightly linked together. Activating one of these representations also reinforces the activation of the other representations. I propose that this system also influences language production and that name-related gazes are caused by the coactivation of all representations.

When speakers look at an object in the apprehension phase a bound representation of this object is created. This representation includes visual, linguistic, conceptual and spatial location representations (see Figure 2.1). In the formulation phase, the linguistic representation is activated which also activates the other features including the spatial location feature. This causes the eyes to move to, or remain at the location of the object. These fixations are the name-related gazes. This mechanism is very similar to the one described by Griffin & Oppenheimer (2006) in language production and by Altmann & Kamide (2007) in language comprehension. However, according to Ferreira, et al. (2008), the system makes some further predictions which were tested in the current experiments.

As just mentioned, the activated spatial location representation causes the eyes to move to the associated location. However, Ferreira, et al. (2008) predicted that by moving to this location, the spatial location representation gets activated even more, which in turn also activates the other representations. Thus, by moving to the location of an object, the linguistic representation also receives more activation which facilitates language production. Furthermore, as shown in language comprehension, people fixate empty regions that were previously occupied by an object when processing this object (Altmann, 2004). This looking at nothing effect is a consequence of the bound representation (Altmann & Kamide, 2007). If name-related gazes are also governed by such a system, the looking at nothing phenomenon should also be expected in language production. Ferreira, et al. (2008) proposed that fixating an empty region previously occupied by an object facilitates processing of object related information. Thus, looking at nothing might also facilitate language production.

In the next section, I give a short overview of the experiments conducted to test these predictions.

3.2 Overview of the experiments

Two experiments were conducted in order to investigate the relationship between visual and linguistic information during language production. The main goal of the experiments was to test how visual and spatial information that are perceived during name-related gazes influence language production. In order to test this question, I manipulated the analogousness of the visual / spatial information and the linguistic information. However, in contrast to the experiment by Griffin & Oppenheimer (2006), I did not change the linguistic information that needed to be produced but the visual/spatial information that was perceived by the speaker.

In order to realize this in an experiment, the visual/spatial features of objects needed to be transformed during the language production process. Thus, participants needed to be exposed to the scene only to perceive its content. In order to accomplish this, the division between the apprehension and the formulation phase in language production was utilised. As mentioned (Section 1.2.2), speakers take about 300-400ms in order to apprehend the content of a scene. Scene perception in the apprehension phase is highly parallel and requires only one fixation (Griffin & Bock, 2000). Thus, to elicit language production, objects in the current experiment were only presented for a period that corresponds to the apprehension phase such that speakers were able to perceive the objects. Following this apprehension phase, location and visual features of objects were changed such that, during the formulation phase, speakers gazed at different objects.

The designs of Experiments 1 and 2 were similar to the experiments by Griffin (2001). In Experiment 1, participants were presented with a set of three objects. The task was to describe the spatial relation of two objects. During the formulation phase, the location, the spatial relation, and the visual features of the objects were changed. Thus, I was able to investigate eye movements towards objects that share only a subset of features with the to-be named objects. Furthermore, the performance of language production was examined. By investigating only trials in which participants fixated the objects, I was able to measure whether fixating an object at a different location or fixating an object with different visual features influences language production.

In Experiment 2, the spatial location feature was examined more closely. Instead of changing the spatial and visual features of the objects, the objects were removed during the formulation phase. I was especially interested whether, during language production, speakers attend to the location in which the objects have been before. Furthermore, was language production facilitated when an empty region associated with a named object was fixated?

3.3 Experiment 1: Looking at objects with varying locations and visual features in language production

3.3.1 Methods

3.3.1.1 Participants

Twenty-four participants took part in Experiment 1. The average age of the participants was 22 (range 19-32). All participants were native speakers of English and had normal or corrected to normal vision. Each participant was paid 5 pounds.

3.3.1.2 Material

Forty-eight items consisting of three objects arranged in a two-by-two grid were used (see Figure 3.1 for an example). A red and a green coloured frame surrounded two objects. These two objects were termed the *critical objects* and were crucial to the production task. The third object was a filler object. The fourth port in the grid remained empty. The objects were selected from the Hemera® Photo-Objects collection or downloaded from Internet sources.

The frequency of the words describing the three objects in one set was kept similar (i.e. the frequency did not vary by more than 0.3 CobSLog (CELEX; COBUILD spoken logarithmic frequency). The number of syllables was kept the same and in most sets the word length differed by not more than three characters. Furthermore, the names of the objects that appeared together in one set were not related in phonological form. Only 36 objects distributed over 12 items were used in the experiment. However, each set of objects was shown four times with varying positions of the objects resulting in forty-eight items. The complete set of the object names including their word frequency is listed in Appendix A.

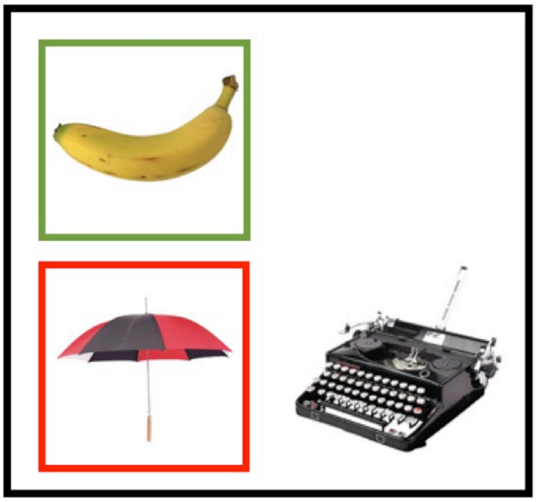

Figure 3.1: An example for a set of objects in Experiment 1. Participants were instructed to describe this set as: *The banana is above the umbrella*.

For each type of object two different tokens were selected. Figure 3.2 illustrates an example of two tokens of the type *umbrella*. The two tokens were selected such that they differed in colour, form and orientation.

Figure 3.2: Two tokens of the object type *umbrella.*

Three research assistants independently named the objects to ensure a consistent naming of objects across participants. Only objects in which at least two of the assistants agreed on one name were included as experimental item. Forty-eight sets of objects were included as fillers.

The sets of objects were presented in the centre of the screen. A set had a size of 278 pixels left to right and 278 pixels top to bottom. Participants sat at a distance of 90cm from the monitor. Thus, the objects were within a visual angle of approximately seven degrees, which is within the parafovea. This was necessary in order to ensure that speakers were able to perceive all three objects with only one fixation.

3.3.1.3 Apparatus

An SR EyeLink 1000 eye tracker with a sampling rate of 1000 Hz and a spatial resolution of less than 1/4 degree was used. The tracker was tower mounted and eye movements of the right eyes were recorded. The chin rest of the eye tracker was removed in order to enable participants to speak. The head was stabilised by using the forehead rest. The stimuli were presented on a 19-inch CRT monitor running at 140Hz and with a resolution of 1024x 768 pixels. To record the speech, a microphone was used. The speech data of each trial was saved in separate wav files. The Experiment Builder software developed by SR research was used to run the experiment. The analysis was done using DataViewer, developed by SR Research Ltd. and Matlab.

3.3.1.4 Procedure

After reading the instructions, participants were seated in a comfortable chair in front of an EyeLink 1000 eye tracker. Before the start of the experiment, participants saw all objects with their respective names. Afterwards, the objects were shown once more and participants had to name them according to the names just shown. According to Wheeldon, Meyer, & Meulen (2007), the repetition of the objects should not influence the speech or the viewing times of the objects. This part of the experiment took approximately 15min. Before calibration, participants were presented with three practice trials after which they could ask questions. The calibration consisted of a nine-point fixation stimulus, which the participants had to fixate one after the other. If the following validation was poor the calibration was repeated.

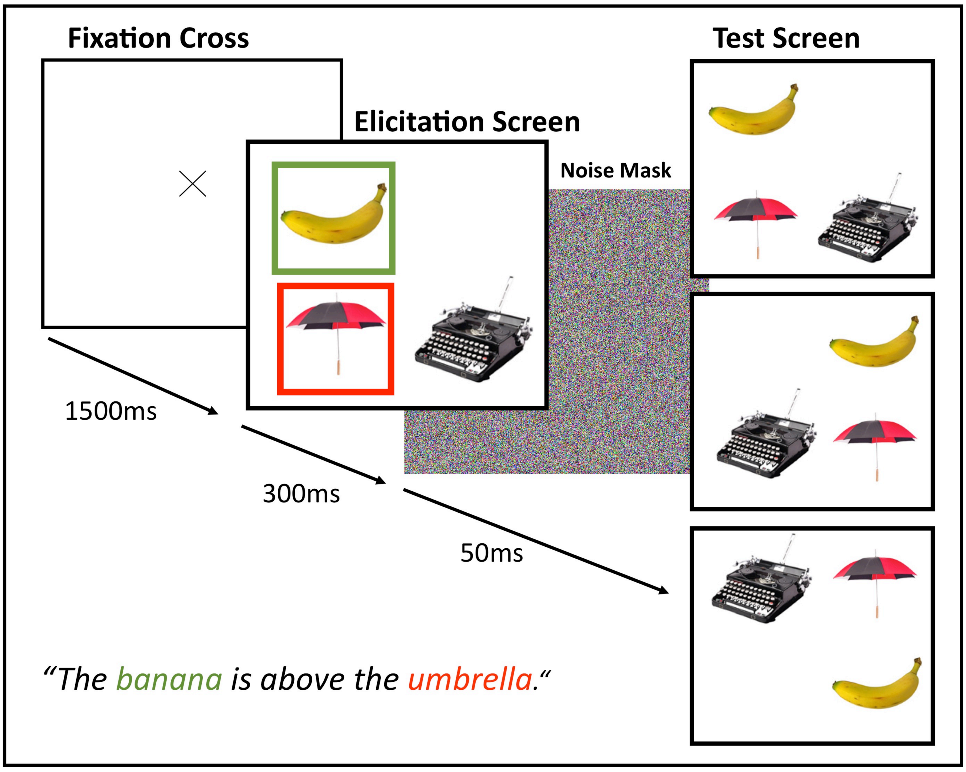

Figure 3.3: Procedure of Experiment 1.

Figure 3.3 shows the procedure of presenting the experimental trials. People saw a fixation cross for 1500ms, which was replaced by the *elicitation screen*. The elicitation screen consisted of three objects. The two critical objects were surrounded by a green and a red frames. The task was to create a sentence by describing the spatial relation between the objects in the green frame and the object in the red frame (see Example (3.1)).

(3.1) *The banana is above the umbrella*.

An object in the green frame had to be named first. The elicitation screen remained visible for 300ms and was then at first replaced by a noise mask, which remained on screen for 50ms, and then by the test screen. The test screen included the objects from the elicitation screen. However, the position of the objects varied between conditions. Furthermore, the coloured frames were not visible in the test screen. For a more detailed description of how objects were arranged refer to the following design section.

People were instructed to start speaking as fast as possible, but since the test screen was visible only 350ms (300ms elicitation screen + 50ms mask) after onset of the elicitation screen, the onset of their speech occurred always after the onset of the test screen. In case, the participants did not see one or both of the objects in the elicitation screen, they were instructed to replace the names of the not recognised objects with the word *something.*

The experiment was divided into four blocks after which participants were asked whether they want a short break. After a break a new calibration was carried out. The set of pictures was repeated in each of the four blocks, but the position of the objects and the frames changed in each block. The pictures were repeated to shorten the practice phase and as a result increase the familiarity of each picture in order to make it easier to recognise them quickly in the elicitation phase of the experiment.

3.3.1.5 Design

A 2 (Token: Same Token, Different Token) by 3 (Position: Same Position. Different Position, Different Relation) design was utilised (see Figure 3.4). In the same token condition, the same token of a type of object was used in the elicitation and in the test screen. In the different token condition, the token differed in elicitation and test screens. Items in the same position conditions were comprised of objects that were located in the same position in both the elicitation and test screens. In the different position condition, the objects changed their position such that the two critical objects moved towards the two remaining slots in the grid (i.e. the slots in which the filler object and the empty slot were located), but without changing their spatial relation to each other. If the banana was above the umbrella in the elicitation screen, it was still above the umbrella in the test screen.

The filler object and the empty slot changed their position in the same manner. In the different relation condition, the two critical objects also moved to the two remaining slots in the grid. However, the spatial relation to each other also changed such that the banana would be below the umbrella. The filler objects and the remaining empty slot moved in the same fashion.

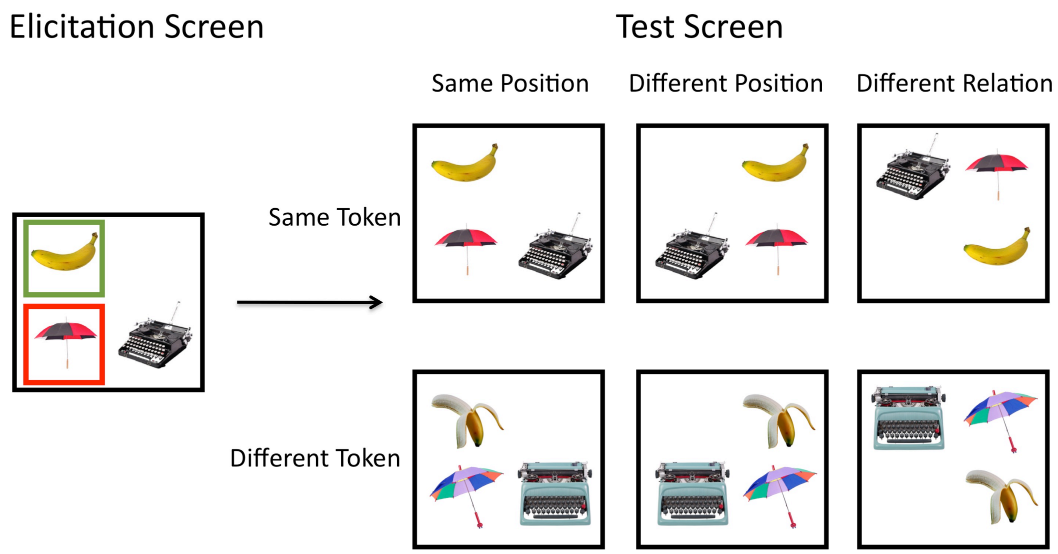

Figure 3.4: Experimental design of Experiment 1.

3.3.1.6 Data Analysis

According to the main question of this study, the analysis was divided into two parts. The first part analysed how speakers directed their eye movements towards objects when naming those objects. Previous studies showed that speakers begin fixating to-be named objects at about 800-1000ms before speech onset and retreat gazes at about 200-100ms before speech onset (see Section 1.2.2). According to these results, I was especially interested in the time frame between 900ms and 100ms before speech onset of a critical noun. This time frame was termed the *critical time frame* and specified either the time before speech onset of the first (e.g. *banana*) or of the second critical object (e.g. *umbrella*).

In the second part of the analysis, I investigated whether speakers benefited from fixating a critical object prior to production even when the object changed its position. An object was labelled as fixated when speakers fixated the object for at least 100ms within the critical time frame. To analyse the benefit from fixating the objects, articulation and speech onset times were measured. Articulation duration was used as a measure of speech disfluency. Less speech disfluency should be reflected by faster articulation times. Speech onset is a measure of the speed of lemma and lexeme selection (Levelt, 1989), and early speech onset indicates fast selection and thus facilitation of language production.

Articulation and speech onset times were measured manually by using the Praat software (Version: 5.0.38; Boersma & Weenick, 1992-2008). Cue points were set in each sound file to label the onsets and offsets of the first and second noun phrase and noun. Spectrograms and the actual sounds were used in order to determine the cue points.

Most of the data of the experiments in this thesis were analysed by using linear mixed models (LME; Baayen, 2008). LME models are particularly useful for comparing groups that do not comprise the same amount of data. The data of the current experiments are unbalanced because items needed to be divided post-hoc into items in which the critical region had been fixated and items in which the critical region was not fixated. Furthermore, only items in which objects were named correctly were used, which created missing data points. In comparison to a repeated measures analysis of variance (ANOVA), an LME model is more robust towards such data. LME models first account for differences in the baseline condition caused by participants and items (random effects) before adding the contribution of the experimental manipulations (fixed-effects).

Figure 3.5 represents a working example of how the results of an LME analysis are visualised. Figure 3.5 shows the mock results of a study investigating effects of frequency and word length on reading times. The mock experiment employed a 2 (Frequency: low frequency, high frequency) by 2 (Word length: short words, long words) design.

The used model is shown in Example **(3.2)**:

```
(3.2) Msat = lmer(RT ~ Frequency * WordLength + (1|S) + (1|I),
            data = file_dat)
```

The model includes a random intercept for each subject and each item. Furthermore, the model tested effects of frequency, word length and the interaction between frequency and word length.

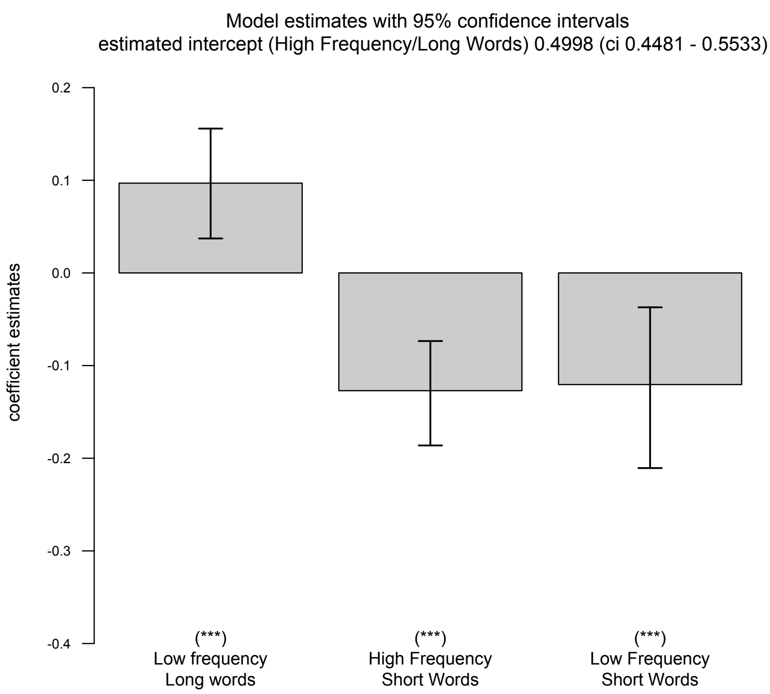

Figure 3.5: Working example to illustrate how such graphs have to be interpreted.

The intercept is presented within the header of the graph in Figure 3.5. The condition high frequency/long words were used as the intercept and the estimated coefficient of the intercept was at 0.4998 with a confidence interval of 0.4481 – 0.5533. This information is very important in order to interpret the meaning of the bars within the graph. The labels in the x-axis show which comparison or interaction a current bar represents. The label below the bar on the left side indicates that the intercept was compared with the condition low frequency and long words. Since the intercept comprised high frequency and long word, differences between high and low frequency words were compared within long words. The bar on the right side represents the frequency × word length interaction. The label below this bar indicates this because both factors (low frequency and short words) differ from the intercept.

The data was not centred. Thus, in contrast to an ANOVA analysis, there are no traditional main effects. Rather, comparisons of a factor only represent a comparison of one half of the data. In the example above, high and low frequencies were only compared with each other within long words. However, the non-significant interaction indicates that the effect is valid for the complete dataset (see Jones, Obregón, Kelly, & Branigan, 2008 for a study that utilised this procedure).

3.3.2 Results

3.3.2.1 Errors

Items were marked as erroneous when participants used an inappropriate word for any of the two objects (e.g. *pen* instead of *banana*), mentioned that they did not see one or both of the objects, used the word *something* to replace the name of an object, used an incorrect verb (e.g. is *below* instead of is *above*), or confused the order of naming (the object in the red frame was named first).

Participants made an error of any of the above kinds in 34.3% of the trials.

Table 3.1 shows the errors for each condition. An ANOVA was carried out to analyse differences between conditions. More errors were made in the condition: different token than in the condition: same token ($F_1(1,23) = 5.89$, $p < .05$; $F_2(1,47) = 13.78$, $p < .001$). Furthermore, a main effect of position was observed ($F_1(2,22) = 7.42$,$p < .01$; $F_2(2,46) = 15.96$, $p < .001$). Pairwise comparisons showed that fewer errors were made in the same position condition than in the different position condition. Furthermore, less error were made in the different position condition than in the different relation condition (all ps < .05). The token × position interaction was not significant ($F_1(2,22) = .81$, p = n.s; $F_2(2,46) = 1.09$, p = n.s).

Most errors were made on the second noun (25.1%) followed by errors on the first noun (15.8%). A wrong verb was used in 11.0% of the trials. The order of the frames was confused in 6.8% of the trials which resulted in a correct but reversed order of naming of the objects (e.g. *The umbrella was below the banana* instead of *The banana was above the umbrella*). A confused order was also coded as an error in all components (i.e. first noun, second noun and verb). The errors of the first and second noun were further analysed.

Table 3.1: Errors made on the first and second noun in Experiment 1

Token	**Same Token**			**Different Token**		
Position	**Same Position**	**Different Position**	**Different Relation**	**Same Position**	**Different Position**	**Different Relation**
All errors	19.3%	29.2%	41.1%	34.9%	34.4%	48.4%
First noun	10.9%	8.9%	18.8%	19.3%	15.6%	22.9%
Second noun	12.5%	18.2%	33.3%	24.5%	28.7%	34.9%

On the first noun, more errors were made in the different token condition than in the same token condition ($F_1(1,23) = 6.63$, $p < .05$; $F_2(1,47) = 8.25$, $p < .01$). Furthermore, the main effect of position was significant ($F_1(2,22) = 4.11$, $p < .05$; $F_2(2,46) = 4.03$, $p < .05$). Pairwise comparisons revealed that fewer errors were made in the different position condition than in the different relation condition (all ps $< .05$). The token × position interaction was not significant ($F_1(2,22) = .27$, p = n.s; $F_2(2,46) = .32$, p = n.s).

The results for the second noun were very similar. A significant main effect of token showed that more errors were made in the different token condition than in the same token condition ($F_1(1,23) = 5.83$, $p < .05$; $F_2(1,47) = 12.49$, $p < .001$). Furthermore, a significant main effect of position was found ($F_1(2,22) = 10.03$, $p < .001$; $F_2(2,46) = 12.43$, $p < .001$). Pairwise comparisons revealed that more errors were made in the different relation condition in comparison to the different position and the same position conditions (all ps $< .05$). The token × position interaction was not significant ($F_1(2,22) = 1.37$, p = n.s; $F_2(2,46) = 1.32$, p = n.s).

3.3.2.2 Language production data

Speech onset and offset times of the sentences, of the first noun phrases, of the first nouns, of the second noun phrases, and of the second nouns are presented in this analysis. The noun phrases consisted either of the determiner *the* and the noun (e.g. *the banana*), or, in case participants omitted the determiner, of the noun only (e.g. *banana)*. Only correct items were used in this analysis. The average onset time of the sentences after the onset of the elicitation screen was 1687ms (Standard deviation (SD): 891ms). The onset time of the sentences also denoted the onset time of the first noun phrases. The mean onset time of the first nouns occurred at 1864ms (SD: 943ms) and the mean offset time of the first noun phrases and first nouns at 2364ms (SD: 987ms) after the onset of the elicitation screen. The mean onset time of the second noun phrases occurred at 3306ms (SD: 1300ms) and the mean onset time of the second nouns at 3449 (SD: 1337ms). The mean offset time of the sentences was at 4000ms (SD: 1360ms). The offset time of the sentences also denoted the offset time of the second noun phrases and of the second nouns.

3.3.2.3 Proportions of fixations overall

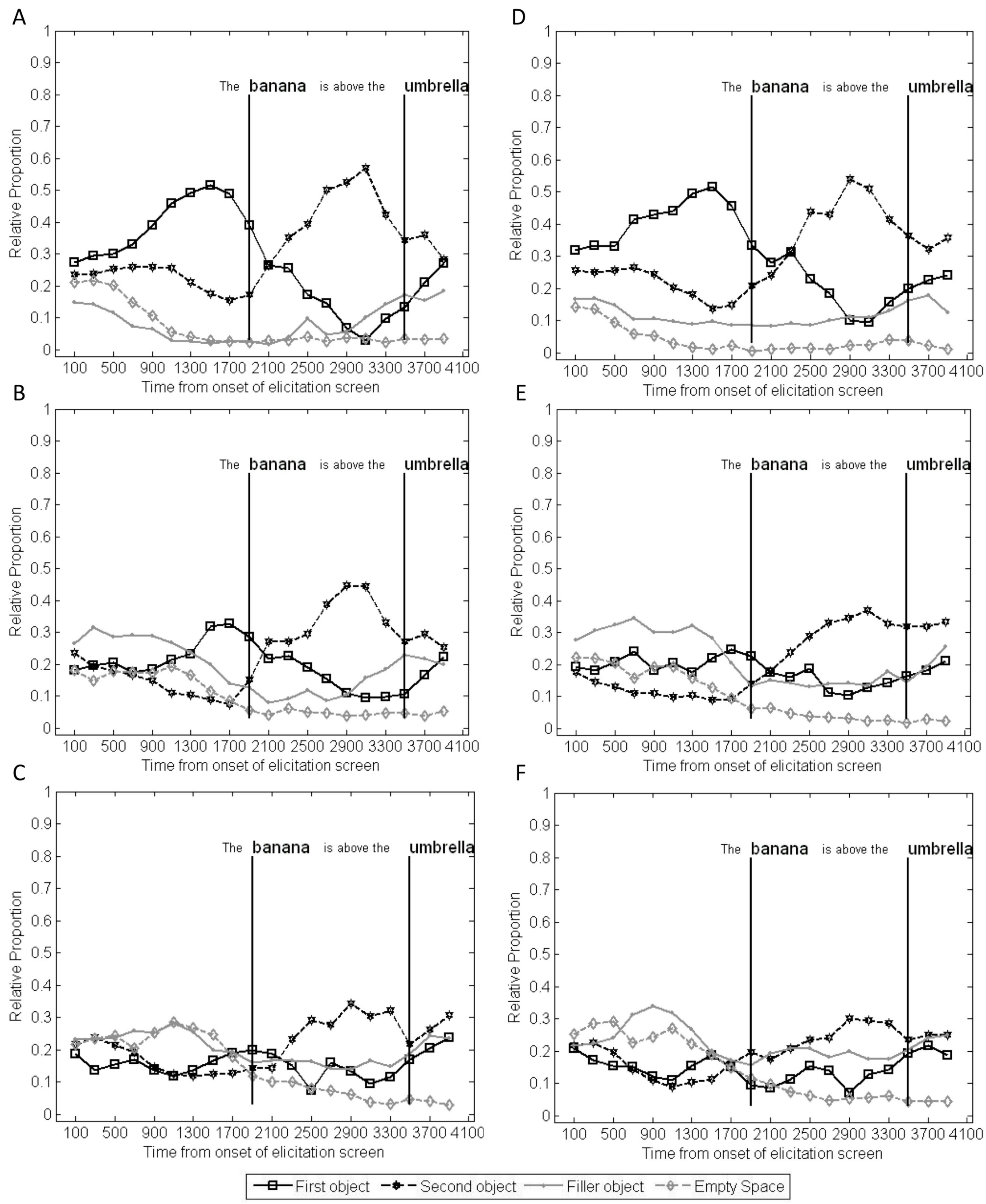

Figure 3.6: Gazes to the objects in the two-by-two grid in relation to the speech onsets of the first and second noun. A: same token & same position, B: same token & different position, C: same token & different relation, D: different token & same position, E: different token & different position, F: different Token & different relation

Before presenting the results of the eye movement data, I first introduce the methodology of synchronising the eye movement and language production data used in the current analysis.

The critical time frame of 900ms to 100ms before speech onset of the nouns was of main interest in the current analysis. By concentrating on this time frame, I was able to include both eye movement and language production data in one plot (see Figure 3.6). Instead of plotting both data on a continuous time line, eye movement data were plotted relative to the speech onsets of the first and the second nouns. For each trial, the speech onset times of the first and the second nouns were measured and eye movements relative to these time points were plotted into the graphs in Figure 3.6. Based on the average speech onset times of the first nouns (1864ms, see Section 3.3.2.2), eye movement data of 1900ms before the speech onset of the first nouns were plotted. Furthermore, 400ms after the speech onset of the first noun were plotted. The average time between onset of the first and second nouns was 1585ms. This time was rounded to 1600ms. The remaining 1200ms were plotted using the speech onset of the second noun as reference. A further 400ms of eye movements occurring after the speech onset of the second noun were plotted.

In summary, this procedure results in an accurate display of the eye movements that occurred 1900ms before and 400ms after the speech onset of the first noun and 1200ms before and 400ms after the speech onset of the second noun. By following this procedure, eye movements within the critical time frame of both the first and second nouns could be depicted in one graph.

In the graphs in Figure 3.6 A-F the x-axes denote the time starting from the onset of the elicitation screen. The y-axes denote the relative proportions of fixations to the three objects, to the empty port in the two-by-two grid and to the area outside the two-by-two grid. However, only fixations within the two-by-two grid are depicted in the graphs. The leftward vertical lines at 1900ms denote the speech onset of the object in the green frame (e.g. *banana*), and the rightward vertical lines at 3500ms denote the speech onset of the object in the red frame (e.g. *umbrella*). Only correct items were used for the following analyses.

I was especially interested whether speakers fixated the first object within the critical time frame of the first noun and the second object within the critical time frame of the second noun. The proportions of fixations towards the three objects and the empty space within the critical time frames were compared with each other. Figure 3.6 A shows the proportions of fixations in the condition: same token & same position. The observed pattern of eye movements was very typical for this kind of language production data. As described in Griffin (2004), when producing a sentence like *The banana is above the umbrella*, speakers gaze at the first to be named object (banana) within the critical time frame. The data of the current experiment confirmed this finding. In the critical time frames before both the onsets of the first and second nouns, speakers fixated more often the first and second objects than other objects or the empty space (all ps < .001). Figure 3.6 B depicts the proportions of fixations for condition: same token & different position. Speakers again fixated the first and second objects before the onsets of the first and second nouns more often than the other objects or the empty space (ps < .01). In condition: same token & different relation (Figure 3.6 C), in the critical time frame, speakers only fixated the second objects more often than other objects before the onset of the second noun (ps < .001). Figure 3.6 D illustrates the proportions of fixations in the condition: different token & same position. In the critical time frame, speakers fixated the first and second objects more often than other objects (ps < .001). In conditions: different token & different position (Figure 3.6 E) and different token & different relation (Figure 3.6 F) participants only fixated the second objects more often than other objects in the critical time frame (ps < .02). The first objects were not fixated more often than other objects within the critical time frame.

To summarise the results, the second objects always received more fixations than the other objects independent from whether the token, the position, or the relation of the objects changed. The results for the fixations on the first objects were not as clear-cut. Only in conditions same token & same position, same token & different position and different token & same position were the first objects fixated more often than other objects within the critical time frame.

In the next section, the relative proportions of fixations towards the first and second objects were compared between conditions.

3.3.2.4 Proportions of fixations during the critical time frame

In this section, the relative proportions of fixations towards the critical objects were compared between conditions. Table 3.2 illustrates the mean proportions within the critical time frame of 900ms to 100ms before speech onset of the critical first and second nouns.

Table 3.2: Proportions of fixations towards the first and second objects during the critical time frame of 900ms to 100ms before the speech onset of the first and second nouns, respectively. Values in parentheses represent standard deviations.

	Same Token			Different Token		
	Same Position	Different Position	Different Relation	Same Position	Different Position	Different Relation
First Object	0.49 (0.02)	0.28 (0.02)	0.16 (0.02)	0.47 (0.02)	0.21 (0.02)	0.15 (0.02)
Second Object	0.50 (0.02)	0.40 (0.03)	0.30 (0.03)	0.47 (0.03)	0.34 (0.03)	0.27 (0.03)

Fixations towards the first object during the critical time frame of the first noun

Speakers fixated the first objects more often when objects did not change their position or relation to each other. Figure 3.7 A shows the output of the LME analysis. Two fixed effects were used: Token (same token, different token) and Position (same position, different position, different relation). Participants and items were included as a random factor. The condition: same token & same position comprised the intercept. To determine whether position had an effect on the dependent variables, the conditions: same token & different condition and same token & different relation were compared to the intercept. In order to test the token fixed effect, the condition: different token & same position was compared to the intercept. The results from the LME analysis are depicted in Figure 3.7.

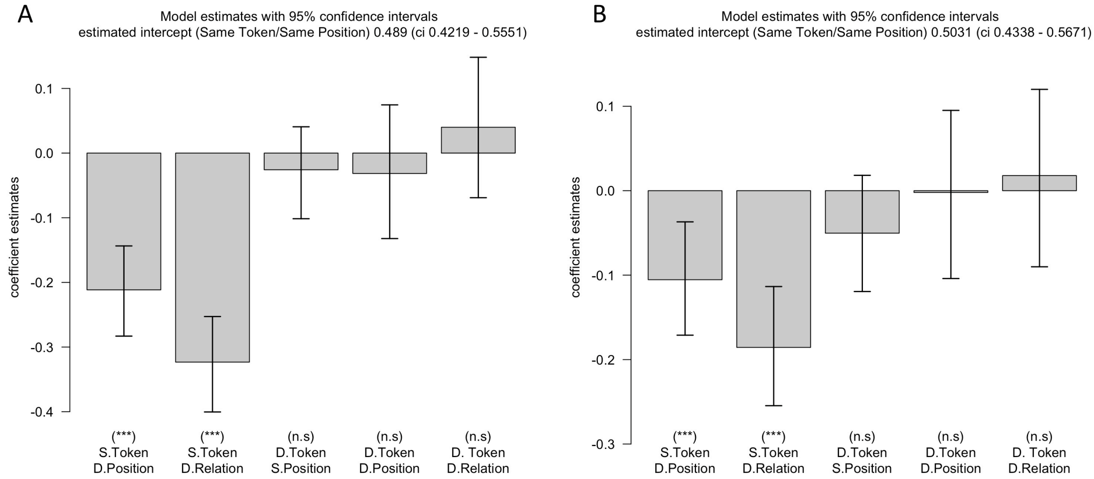

Figure 3.7 A: Analysis of the fixations towards the first objects within the critical time frame of the first nouns. B: Analysis of fixations towards the second objects within the critical time frame of the second nouns. S abbreviates same and D abbreviates different.

Participants showed significantly fewer fixations on the first objects in the same token & different position condition ($t = 6.02$, $p < .01$) and in the same token & different relation condition ($t = 8.67$, $p < .01$) in comparison to the intercept. No significant difference between the condition: different token & same position and the intercept ($t = .73$, $p = .47$) was measured. Furthermore, all interactions were not significant (different token & different position: $t = .60$, $p = .55$; different token & different relation: $t = .76$, $p = .45$).Fixations towards the second object during the critical time frame of the second noun

The proportions of fixations towards the second object were almost identical with the results from the first object (Figure 3.7 B). Significantly fewer fixations were measured on the second objects when the position of the objects changed ($t = 3.11$, $p < .01$) as well as when the relation of the objects changed ($t = 5.15$, $p < .01$). The different token & same position condition was not significantly different from the intercept ($t = 1.45$, $p = .15$). Furthermore none of the interactions were significant (different token & different position: $t = .03$, $p = .98$; different token & different relation: $t = .36$, $p = .72$).

3.3.2.5 Articulation time and speech onset time

In this section, the articulation times and speech onset times of the critical nouns were compared between conditions. Only items were included in which participants fixated the critical objects within the critical time frames before onset of the first and second nouns. Table 3.3 lists the articulation times of the first and second nouns and noun phrases. The condition: same token & same position comprised the intercept.

Table 3.3: Articulation times of the first NP and noun and of the second NP and noun for items in which the respective objects of the nouns were fixated. Values in parentheses represent standard deviations.

	Same Token			**Different Token**		
	Same Position	**Different Position**	**Different Relation**	**Same Position**	**Different Position**	**Different Relation**
First NP	668 (28)	691 (50)	672 (43)	642 (26)	772 (58)	649 (94)
First Noun	482(14)	502 (19)	482 (19)	477 (14)	518 (28)	463 (29)
Second NP	692 (21)	691 (19)	678 (20)	660 (20)	747 (53)	748 (33)
Second Noun	562 (13)	565 (13)	555 (13)	547 (13)	530 (19)	571 (19)

Articulation times of the first and second nouns and noun phrases

Figure 3.8 A and Figure 3.8 B show the outputs of the LME analyses for the articulation times of the first noun phrase and noun respectively. No differences between conditions were found in either the first noun or the first noun phrase.

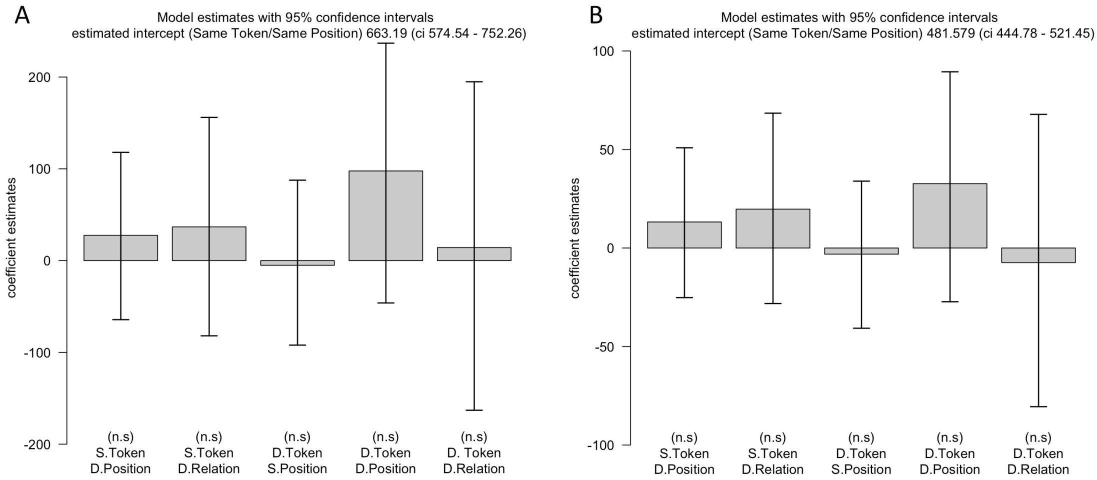

Figure 3.8 A: Articulation times of the first noun phrase. B: Articulation times of the first noun.

Similar results were found for articulation times for the second noun phrase and the second noun. As shown in Figure 3.9 A and Figure 3.9 B, none of the comparisons were significant.

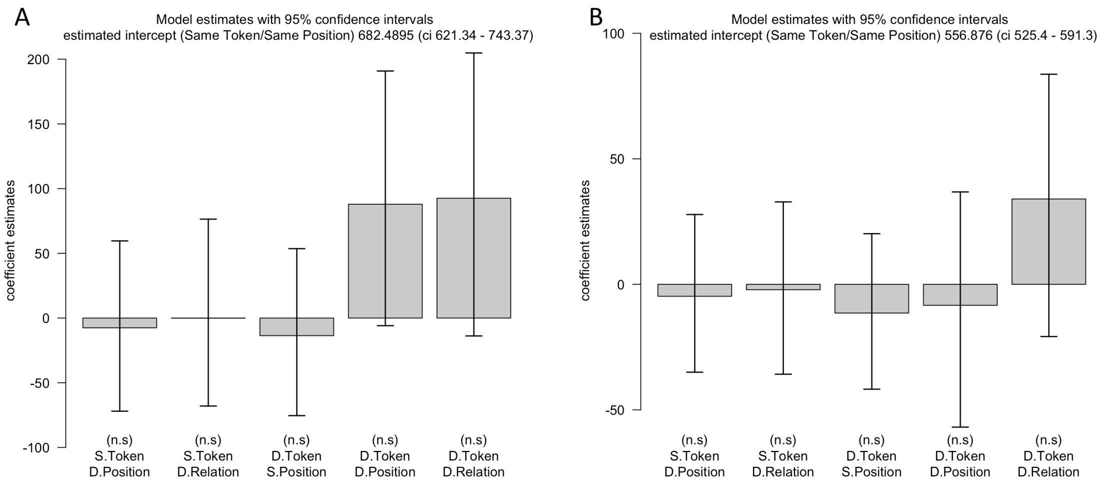

Figure 3.9 A: Articulation times of the second noun phrase. B: Articulation times of the second noun.

Speech onset times of the first and second noun phrases and nouns

The speech onset times of the first and second noun phrases and nouns are depicted in Table 3.4. Speech onset times were measured starting from the onset of the elicitation screen. Figure 3.10 shows the outputs of the LME analysis for speech onset times of the first noun phrase and noun.

Speech onset for the first noun phrase (Figure 3.10 A) was slower in the same token & different position condition in comparison to the intercept ($t = 2.8$, $p < .01$). Furthermore, speech onset was slower in the different token & same position condition in comparison to the intercept ($t = 2.8$, $p < .01$). A significant interaction (different token & different position; $t = 2.6$, $p < .01$) was found. Pairwise comparisons revealed that speech onset in the different token & different position condition did not differ from the different token & same position ($t = .5$) or from the different position & different relation condition ($t = .9$). Thus the manipulation of the position of objects affected speech onset times only when the same tokens were shown in the test screen. Furthermore, the different token & different position condition did not differ from the same token & different position condition ($t = 1.1$). Thus, the token manipulation had different effects when the position of the objects changed. Speech onset times were only slower when the token changed and objects did not change their positions. However, speech onset times did not change when the token changed and the position differed.

Table 3.4: Speech onset times of the first NP and noun and of the second NP and noun for items in which the respective objects of the nouns were fixated. Values in parentheses represent standard deviations.

	Same Token			Different Token		
	Same Position	**Different Position**	**Different Relation**	**Same Position**	**Different Position**	**Different Relation**
First NP	1618 (65)	2029 (126)	1869 (149)	1866 (117)	1845 (117)	1912 (185)
First Noun	1805 (69)	2218 (130)	2059 (153)	2031 (119)	2099 (133)	2098 (231)
Second NP	3143 (93)	3561 (158)	3270 (137)	3392 (125)	3444 (142)	3426 (186)
Second Noun	3273 (97)	3688 (160)	3392 (136)	3506 (126)	3661 (150)	3603 (199)

Figure 3.10 B illustrates the speech onset times of the first noun. The results were very similar to the speech onset times of the first noun phrase. Speech onset time was slower in the same token & different position condition in comparison to the intercept ($t = 2.7$, $p < .01$). Furthermore, speech onset time was slower in the different token & same position in comparison to the intercept ($t = 2.7$, $p < .01$). A significant interaction (different token & different position; $t = 2.2$, $p < .05$) was found. Pairwise comparisons showed no significant differences in the position manipulation within the different token conditions (all ts < .3). Thus, as in the analysis of the first noun phrase, the manipulation of the position of the objects had an effect only when the token remained the same in the elicitation and the test screen. Furthermore, the different token & different position condition did not differ from the same token & different position condition ($t = .7$). Speech onset times were slower when the token changed but object did not change its position. However, speech onset times did not change when the token changed but position differed.

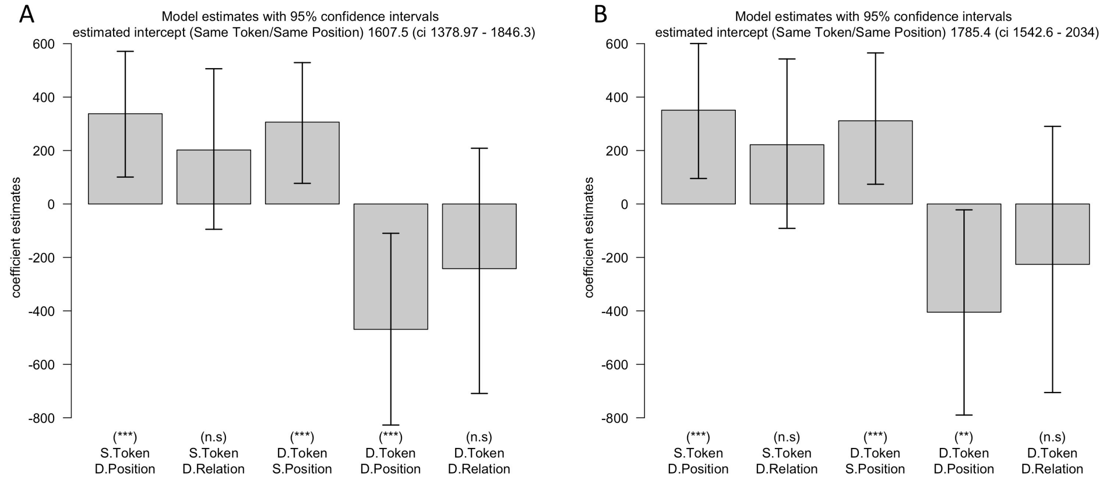

Figure 3.10 A: Speech onset times of the first noun phrase. B: Speech onset times of the first noun.

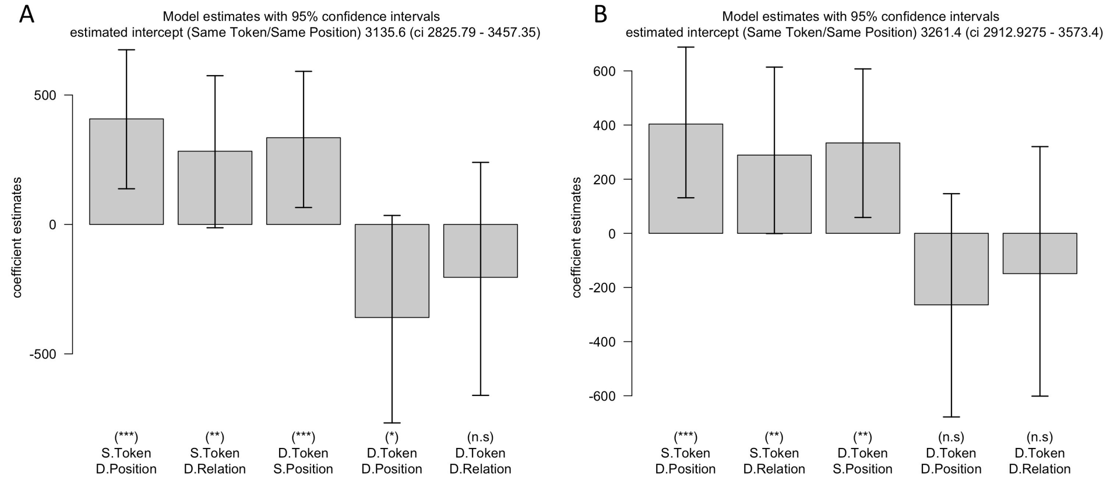

Figure 3.11: A: Speech onset times of the second noun phrase. B: Speech onset times of the second noun.

Figure 3.11 illustrates the speech onset times of the second noun phrase (Figure 3.11 A) and the second noun (Figure 3.11 B). In comparison to the intercept, the speech onset of the second noun phrase was slower in the same token & different position condition ($t = 3.0$, $p < .01$) and in the same token & different relation condition ($t = 2.0$, $p < .05$). Furthermore, speech onset time was slower in the different token & same position condition in comparison to the intercept ($t = 2.6$, $p < .01$). None of the interactions were significant ($ts < 1.8$).

Very similar results were found for the onset times of the second noun. Speech onset time was slower in the same token & different position condition in comparison to the intercept ($t = 2.9$, $p < .01$). Likewise, the same token & different relation condition was slower than the intercept ($t = 2.0$, $p < .05$). Furthermore, speech onset was slower in the different token & same position condition in comparison to the intercept ($t = 2.5$, $p < .05$). None of the interactions were significant (ts < 1.3).

3.3.3 Discussion

In Experiment 1, the role of spatial location and visual information was tested on performances in speech production. The positions of to-be named objects were changed during the language production formulation phase which occurred 300-400ms after onset of a scene (Griffin & Bock, 2000). Furthermore, the tokens of objects were changed in order to investigate effects on language performance caused by a variation of the visual features.

At first, name-related gazes were tested when the position and the token identity of objects were changed. Speakers were less likely to fixate to-be named objects when they changed the position in comparison to when the position of the objects remained the same. Changing the relation and the position of the objects affected name-related gazes even more. Speakers looked even less at to-be named objects. Manipulation of the token identity of objects had no effect on name-related gazes. However, in most of the conditions, speakers fixated to-be named objects more often than other objects or the empty space within the two-by-two grid. Especially the second critical object was still fixated more often than the other objects or the empty space.

In general, name-related gazes were very similar to gazes in earlier reports (Griffin, 2004b; Meyer, et al., 1998). Speakers fixated the to-be named objects at a time frame before the speech onsets of the associated nouns. However, in the current study, there was a clear decline of name-related gazes depending on the agreement of spatial relation features between objects in the elicitation and test phase. Varying visual features of the objects did not alter name-related gazes.

These results support the referential hypothesis (Griffin & Oppenheimer, 2006). Only objects that were located at their original positions attracted name-related gazes. Thus, the spatial index caused the eyes to move or remain at the position of the to-be named objects during the formulation phase. Furthermore, differing visual features did not affect name-related gazes as would have been predicted by the content hypothesis.

In order to test how name-related gazes affect language production, the articulation times and speech onset times of the critical noun phrases and nouns were tested. Only items were included in the analysis in which the critical objects were fixated during the critical time frames. Thus, it could be tested how varying spatial location and visual information influences language production. Articulation time, which was used as a measure for disfluency, was not affected by varying spatial location or visual information. However, significant effects of speech onset times were found. When the position of objects was changed, speech onset times of the first and second noun phrases and nouns occurred later in comparison to when the position of the objects did not change. Furthermore, an effect of token was found. When the visual features of the critical objects changed between the elicitation and the test phases, speakers named objects later in comparison to when they did not change. These results are evidence for the theory proposed by Ferreira, et al. (2008). Reactivating the spatial location or visual features facilitates language production.

To summarise, name-related gazes are mostly driven by the spatial location representation of the objects. If objects change their position, speakers are less likely to fixate to-be named objects. Varying visual features did not change the viewing behaviour. This supports the objects file theory, suggesting that the spatial location representation takes a prominent role in retrieving the memory representations (Kahneman, et al., 1992). Furthermore, language production is facilitated when speakers fixate an object at its original position in comparison to the same object at a different position. However, if objects change their positions and speakers fixate the objects, complying visual information of the objects also facilitate language production.

In the following experiment, the role of the spatial location representation is investigated in more detail. In Experiment 1, manipulating the spatial location feature was accomplished by changing the position of an object. However, this methodology is unsuitable to investigate the role of spatial location in isolation. When people looked at a correct or incorrect position, they always also perceived visual information which might interact with the spatial location information. Thus in Experiment 2, objects were not shown in the test phase and I utilised the looking at nothing effect found in language comprehension.

3.4 Experiment 2: Looking at nothing during language production

According to the model by Ferreira, et al. (2008), activating the linguistic representation of an object activates its spatial location representation which causes the speaker to fixate the location of the object even if it is not present anymore. Furthermore, looking at such an empty space leads to a stronger activation of the spatial location representation which should cause a stronger activation of the linguistic representation. Thus, by looking at empty space, language production should be facilitated. This was tested in Experiment 2. I also tested whether speakers fixate the location, at which an object was previously located, more often when the name of an object was harder to retrieve. Previous studies found that name-related gazes towards to-be named objects were longer when objects were low codable in comparison to high codable objects (Griffin, 2001). As introduced in Section (1.2.1), codability affects retrieval of a word's lemma. In the present experiment, it was tested whether this effect can be also observed for gazes towards locations in which an object was previously located.

The methodology of Experiment 2 slightly differed from the methodology of Experiment 1. Speakers saw the set of objects in the elicitation screen. However, objects did not reappear in the test screen. Furthermore, a relatively high error rate was observed in Experiment 1. This might have been caused by a too brief presentation of the objects. Thus, the exposure time of the elicitation screen was extended by 200ms.

3.4.1 Methods

3.4.1.1 Participants

Twenty-four participants took part in Experiment 2. The average age of the participants was 20.8 (range: 18-32). All participants were native speakers of English and had normal or corrected to normal vision. Each participant was paid four pounds.

3.4.1.2 Material

Forty-eight items consisting of a set of three objects were used. The objects were selected from the Hemera® Photo-Objects collection or downloaded from Internet sources. The difficulty of the objects was manipulated by varying the codability of the objects (for a description of codability refer to Section 1.2.1). Codability was tested in a pretest.

Twenty-four participants took part in this web-based pretest using the WebExp2 programme (Keller, Gunasekharan, Mayo, & Corley, 2009). 337 objects were presented consecutively and people were instructed to type in the first name that came to mind. The names, the times of onset of writing and the response times were recorded. In order to measure codability, the H-value was calculated (Snodgrass & Vanderwart, 1980). The H-value was computed by using the Equation 3.1, where k is the number of names given to an object and p_i is the proportion of subjects given one specific name.

Equation 3.1: $$H = \sum_{i=1}^{k} p_i \log_2\left(\frac{1}{p_i}\right)$$

An H-value of 0 denotes 100% of agreement (i.e. only one word was used to name an object) and thus high codability. The higher the H-value, the smaller agreement and the lower codability were. The advantage of the H-value over just using the proportion of the most frequent word is that the H-value takes the number of different names into account. For example, an object that is named in 80% by name A, in 10% by name B, and in 10% by name C has a lower H-value than an object that is named in 80% by name A and in 20% by name B. The proportions of the most frequent words are, however, the same. To divide the objects into high and low codability, 72 objects with the highest and 72 objects with the lowest values were selected. The codability H-values for each of the objects are listed in Appendix B.

The same selection criteria as in Experiment 1 were used to assign objects together in one set. Twenty-four sets of objects were included as fillers.

3.4.1.3 Apparatus

The same apparatus as in Experiment 1 was used.

3.4.1.4 Procedure

The procedure was similar to the procedure in Experiment 1.

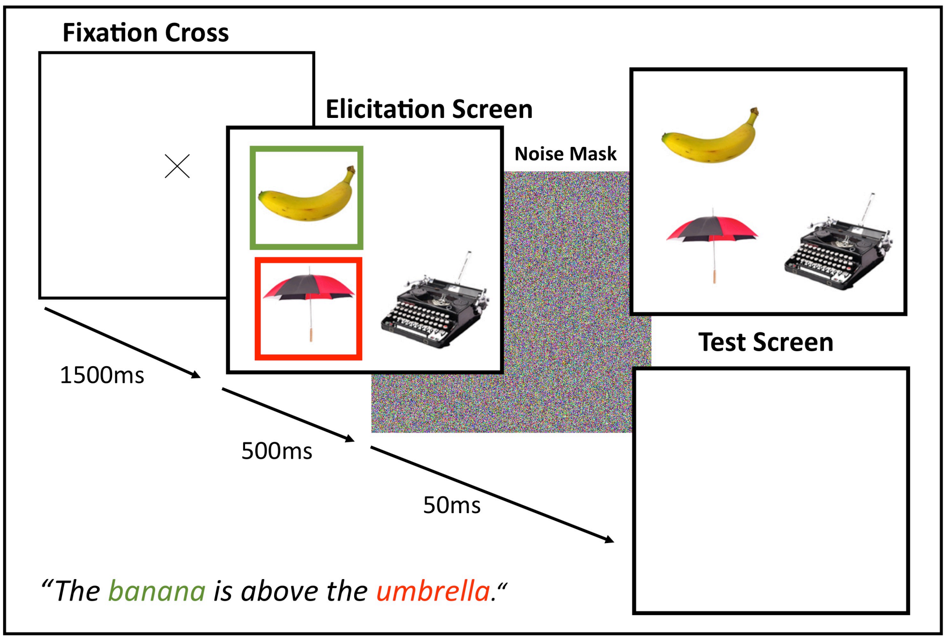

Figure 3.12: Procedure of Experiment 2.

However, objects in the elicitation screen were visible for 500ms. The time was increased in order to make recognition of the objects easier and thus generate fewer errors. Furthermore, the test screen was either completely empty or the three objects from the elicitation screen appeared in the same location in which they were in the elicitation screen (see Figure 3.12). For a more detailed description of how objects were arranged refer to the following design section.

3.4.1.5 Design

The experiment employed a 2 (Visibility: visible, empty) by 2 (Codability: high codable, low codable) design (see Figure 3.13).

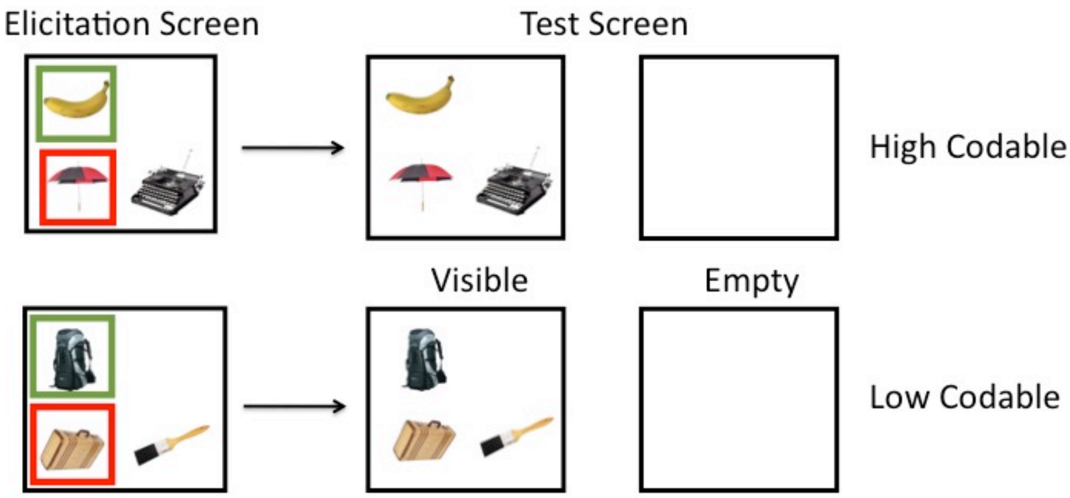

Figure 3.13: Experimental design of Experiment 2.

In the visible conditions, the objects used in the elicitation screen appeared again in the test screen. The arrangement of the objects was exactly the same except that the coloured frames were not visible. In the empty conditions the screen was completely blank. In the high and low codable conditions the codability of the three objects in one set of objects was either high or low respectively.

3.4.1.6 Data Analysis

The data analysis was identical to the analysis in Experiment 1. The analysis was again divided into two parts. The first part analysed how speakers directed their eye movements towards objects and the empty region in which an object has been when naming the objects. In the second part of the analysis, I investigated whether speakers benefited from fixating a critical object or the empty region in which a critical object has been.

3.4.2 Results

3.4.2.1 Errors

The same criteria from Experiment 1 were used to mark items as erroneous. In total, participants made an error in 31.8% of all trials. The number of errors varied relatively strongly between conditions. Table 3.5 summarises the errors. When the objects remained on the screen (visible), 9.7% of the items were named incorrectly in the high codable condition and 24.7% in the low codable condition. When the objects did not remain on the screen (empty) 33.7% and 59.0% of the items where incorrect in the high and low codable conditions respectively.

Table 3.5: Errors made on the first and second nouns in Experiment 2

Visibility	Visible		Empty		All
Codability	High	Low	High	Low	
All errors	9.7%	24.7%	33.7%	59.0%	31.8%
First noun	7.6%	16.3%	11.1%	27.1%	15.5%
Second noun	9.0%	20.5%	28.1%	50.7%	27.1%

An ANOVA analysis revealed significant main effects of codability ($F_1(1,23) = 77.80, p < .001$; $F_2(1,92) = 22.19, p < .001$)[2]. Fewer errors were made in the high codable condition than in the low codable condition. Furthermore, a significant main effect of visibility ($F_1(1,23) = 79.35, p < .001$; $F_2(1,92) = 46.55, p < .001$) was found. Participants made fewer errors in the visible than in the empty condition. The interaction of codability × visibility was only significant in

[2] Since codability was a between item factor, the per items analysis was conducted by using an univariate analysis of variance.

the per subjects analysis ($F_1(1, 23) = 5.52$, $p < .05$; $F_2(1,92) = 1.48$, p = n.s). Most errors were made on the second noun (27.1%), followed by errors on the first noun (15.5%). In 8.4% of the trials the verb was used wrongly and in 6.5% the participants confused the order of naming which resulted in a correct but reversed naming of the objects. A confused order was coded as an error in all components (i.e. first noun, second noun and verb). The number of errors made on the first or second noun varied between conditions. A significant main effect of visibility was found in naming both the first and the second nouns (First noun: ($F_1(1,23) = 11.76$, $p < .01$; $F_2(1,92) = 7.58$, $p < 0.01$); Second noun: $F_1(1,23) = 72.15$, $p < .001$; $F_2(1,92) = 40.66$,$p < .001$). Furthermore a main effect of codability was found (First noun: $F_1(1,23) = 27.93$,$p < .001$; $F_2(1,92) = 22.72$, $p < .001$; Second noun: $F_1(1,23) = 100.22$, $p < .001$; $F_2(1,92) = 19.37$,$p < .001$). The Visibility × Codability interaction was marginally significant for the first noun in the per subjects analysis but not significant in the per items analysis ($F_1(1,23) = 3.50$, $p < .10$; $F_2(1,92) = 1.98$, p = n.s). For the second noun the interaction was significant in the per subjects analysis but not significant in the per items analysis ($F_1(1, 23) = 5.66$, $p < .05$; $F_2(1,92) = 2.07$, p = n.s).

Errors were also compared between trials in which participants fixated the critical position before naming the associated object and trials in which they did not fixate the critical position. However, only trials were tested in which the objects did not appear again. To test whether looking at the empty region had an effect on the error rate, the items were post-hoc divided into items in which people looked at the critical position and items in which they did not (for how this division was carried out see Section 3.4.1.6). Errors of the first noun were tested when subjects fixated and when they did not fixate the first object. Errors of the second noun were tested when people fixated and when they did not fixate the second object. Table 3.6 illustrates the error rates. A mixed logit model (Jaeger, 2008), which is a generalized linear mixed model for binomially distributed dependent variables was employed in order to test for differences.

Table 3.6: Comparison of errors between items in which the critical empty position has been fixated in comparison to when it was not fixated. Only items of condition empty were used.

Fixated	**One or more fixations**		**No fixation**		**All**
Codability	**High**	**Low**	**High**	**Low**	
First noun	10.9%	22.5%	11.8%	35.6%	20.2%
Second noun	28.7%	50.8%	26.3%	50.5%	39.1%

Two fixed effects were utilised: Fixated (one or more fixations, no fixation) and Codability (high codable, low codable). The condition one or more fixations & high codable comprised the intercept. The only significant effects for both the first and second noun were effects between the one or more fixations & low codable condition and the intercept (first noun: $z = 3.22$, $p < .01$; second noun: $z = 3.07$, $p < .01$). This result indicated that more errors were made when the objects were low codable in comparison to high codable objects. However, whether speakers fixated the location in which an object had been or did not fixate this region had no effect on the error rate.

3.4.2.2 Language production data

In the following section the language production data is summarised. Only correct items were used in this analysis. The average speech onset time of the sentences was 1394ms (SD: 612ms) after the onset of the elicitation screen. The onset time of the sentences also denoted the onset time of the first noun phrases. The mean onset time of the first nouns occurred at 1629ms (SD: 795ms) and the mean onset time of the second nouns occurred at 3056ms (SD: 1131ms) after the onset of the elicitation screen. The mean offset time of the sentences was at 3590ms (SD: 1169ms). The offset time of the sentences also denoted the offset time of the second nouns. The offset time of the first nouns and the onset time of the second NPs were also measured. However, since I was especially interested in the empty screen conditions, and due to time constraints, these onset times were only collected for the data which included the empty screen conditions. The mean offset time of the first nouns occurred at 1961ms (SD: 581ms) and the mean onset time of the second NPs at 2734ms (SD: 815ms) after the onset time of the elicitation screen.

3.4.2.3 Proportions of fixation overall – Do speakers look at empty regions

In order to illustrate the fixations within the two-by-two grid, I used the same procedure as in Experiment 1 (Figure 3.6). Onset times of the first and second noun differed slightly and thus the graphs were adjusted accordingly. The mean speech onset time of the first noun was at 1629ms. Thus 1600ms prior to speech onset of the first nouns were plotted. The average time between onset of the first and second nouns was 1427ms which was rounded to 1400ms in the graphs. The first 400ms were plotted based on the speech onset of the first noun and the remaining 1000ms were plotted based on the speech onset of the second noun.

Only fixations within the two-by-two grid are depicted in the graphs. The leftward vertical lines at 1600ms denote the average speech onset time of the object in the green frame (e.g. *banana*) and the rightward vertical lines at 3000ms denote the average speech onset time of the object in the red frame (e.g. *umbrella*). Only correct items are used for the following analyses. Figure 3.14 A shows the fixations within the two-by-two grid for the condition: high codability & visible. T-tests revealed that within the critical time frame before speech onset of the first noun, speakers fixated the first object significantly more often than the other objects or the empty space (ps < .01). Before speech onset of the first noun, speakers already started to fixate the second noun (umbrella). Within the critical time frame of the second object, speakers fixated the second object more often than the other objects (ps < .01).

Figure 3.14 B illustrates the proportions of fixations within the two-by-two grid in the condition: high codable & empty. In this condition, the objects did not reappear in the test screen. Thus, speakers were looking at an empty screen during language production.

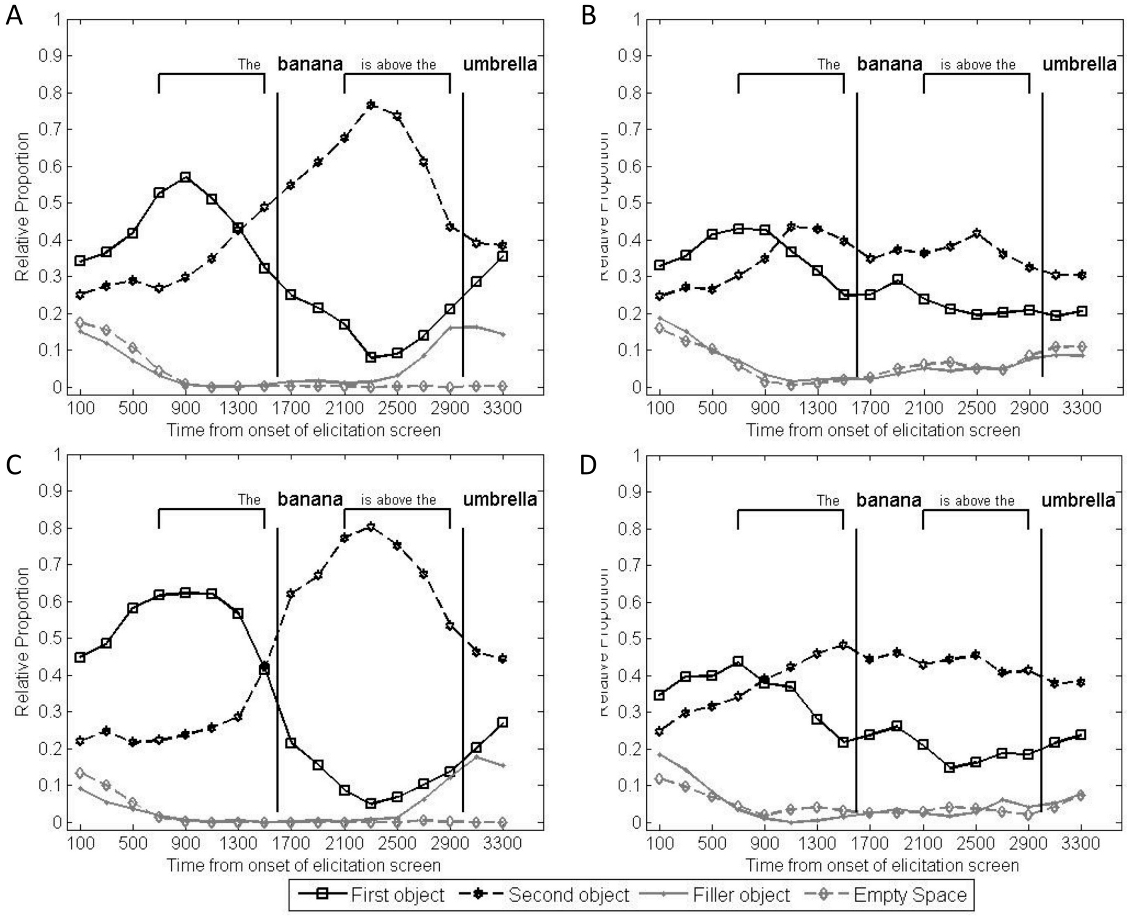

Figure 3.14: Gazes to the objects in the two-by-two grid in relation to the speech onsets of the first and second nouns. A: high codable & visible; B: high codable & empty; C: low codable & visible; D: low codable & empty

Similar to the high codable & visible condition, speakers started to fixate the position in which the first object had been at about 1000ms before speech onset of the first noun. However, they started to move their eyes towards the position in which the second object has been earlier than in the high codable & visible condition. Within the critical time frame, speakers did not fixate the first object more often than the second object ($t = .90$, $p = .36$). However, most importantly, speakers fixated the first object more often than the filler object or the empty space ($ps < .01$). Furthermore, speakers fixated the second objects more often than the other objects or the empty screen within the critical time frame before speech onset of the second noun ($ps < .01$).

Figure 3.14 C illustrates the distribution of fixations in condition: low codable & empty. Within the critical time frame before onset of the first and the second noun, speakers were more likely to fixate the first and second object than the other objects or the empty space (ps < .01).

Likewise, in the condition: low codable & empty (Figure 3.14 D), fixations within the critical time frames of the first and second objects were more likely to occur at the positions in which the first and second object have been, respectively (ps < .05).

3.4.2.4 Proportions of fixations during the critical time frame

In this section, fixations within the critical time frame were compared between visible and empty objects in the test screen and between high and low codable objects. Of main interest was whether people fixated empty critical regions more often when the to-be produced object was low codable and thus more difficult to retrieve.

Table 3.7 shows the proportions of the fixations within the critical time frame before the onset of the first and the second noun for all four conditions. Only correct items were used in this analysis resulting in an unbalanced number of items in the different conditions. Thus, to analyse the data, an LME model was utilised.

Table 3.7: Proportions of fixations towards the first and second objects during the time frame of 900ms to 100ms before the speech onset of the first and second nouns, respectively. Values in parentheses represent standard deviations.

Visibility	**Visible**		**Empty**	
Codability	**High**	**Low**	**High**	**Low**
First Object	0.50 (0.02)	0.59 (0.02)	0.37 (0.02)	0.34 (0.03)
Second Object	0.70 (0.02)	0.76 (0.03)	0.38 (0.02)	0.44 (0.04)

Fixations towards the first object during the critical time frame of the first noun

Two fixed effects were used: Codability (High, Low) and Visibility (Visible, Empty). Participants and items were included as random factors. The condition high codable & visible comprised the intercept. Figure 3.15 A illustrates the LME analysis for the fixations on the first object. Participants showed significantly more fixations on the first objects in the low codable & visible condition (t = 3.19, p < .01) in comparison to the intercept. Furthermore, in the high codable & empty condition, the first object was significantly less often fixated than in the intercept condition (t = 4.41, p < .001). The interaction between difficulty and visibility was significant (t = 2.71, p < .01). These results indicate that low codable objects were fixated more often when objects were visible in the test screen. Furthermore, high codable objects were fixated more often in the visible in comparison to the empty condition. The significant interaction indicates that the low codable & empty items were fixated significantly less often than expected on the basis of the differences between high and low codable in the visible condition. A pairwise comparison revealed no significant difference of high versus low codable items in the empty condition (t = .67).

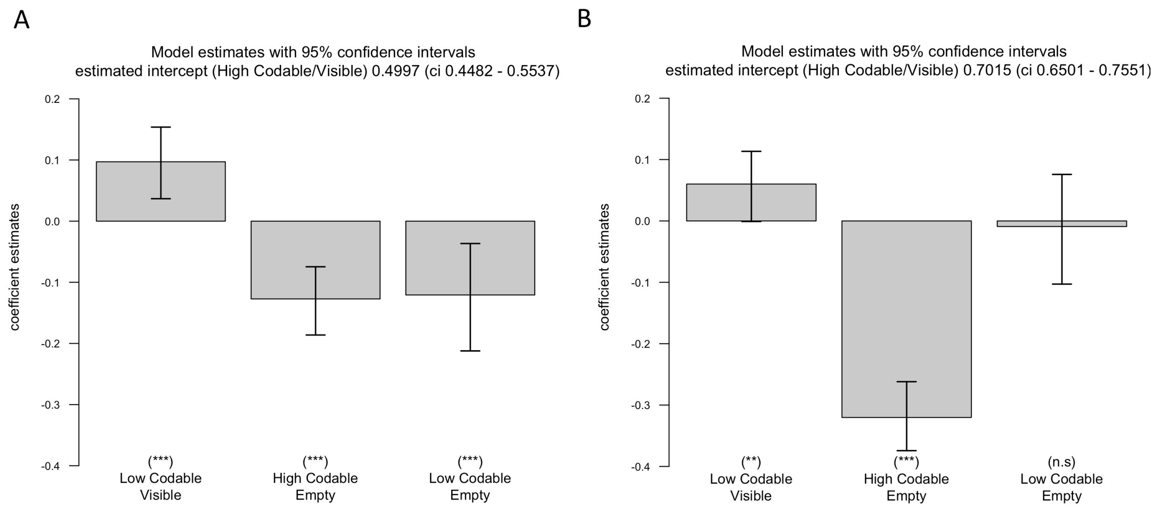

Figure 3.15. A: Analysis of the fixations toward the first objects within the critical time frame of the first noun. B: Analysis of fixations towards the second objects within the critical time frame of the second noun.

Fixations towards the second object during the critical time frame of the second noun

Figure 3.15 B illustrates the output of the LME analyses of fixations on the second object within the critical time frame before the onset of the second noun. Participants fixated the second object in the low codable & visible condition more often than the intercept condition ($t = 2.02, p < .05$). Furthermore, participants fixated the second object in the high codable & empty condition less often than in the intercept condition ($t = 11.03, p < .001$). No significant interaction ($t = 0.24, p = .81$) was found. Most interestingly, the non-significant interaction indicates that in the empty test screen condition, people fixated the position in which the second object has been more often when the second object was low codable in comparison to when it was high codable.

3.4.2.5 Articulation time and speech onset time

To evaluate whether fixating the region in which an object had been facilitates speech production performance, the articulation and the speech onset times of the first and second critical nouns and noun phrases were measured. Only conditions in which the objects did not appear again (Visibility: empty) were used in this analysis. Items in which participants fixated the regions where the first or second objects had been (one or more fixation) were compared with items in which the critical empty regions were not fixated (no fixations). An object was labelled as fixated when the participants fixated it within the critical time frame between 900ms and 100ms before speech onset of the first or second object. This division of the items was made post-hoc.

LME analyses were used to analyse the data. Two fixed effects were utilised: Codability (high codable, low codable) and Fixated (No fixation, One or more fixations). The condition high codable & one or more fixations comprised the intercept.

If fixating the empty position in which objects had been facilitates language production, faster articulation duration and faster speech onset times should be observed in the one or more fixations condition in comparison to the no fixation condition. Furthermore, according to the previous analysis (Section 3.4.2.4), participants fixated an empty region more often when it was associated with a low codable in comparison to a high codable object. Low codable objects are generally more difficult to produce than high codable objects. If fixating empty regions facilitates language production, I expected an interaction between codability and fixated such that the difference in articulation time and speech onset time between high and low codable is smaller when speakers fixated the empty region.

Articulation times of the first and second nouns and noun phrases

Table 3.8 illustrates the articulation times of the first and second critical noun and noun phrases.

Table 3.8: Articulation times of the first noun phrase and noun and of the second NP and noun for items in which the respective objects of the nouns were fixated. Values in parentheses represent standard deviations.

Fixated	**One or more fixations**		**No Fixations**	
Codability	**High**	**Low**	**High**	**Low**
First NP	562ms (20)	663ms (32)	623ms (27)	933ms (79)
First Noun	413ms (10)	486ms (18)	432ms (16)	536ms (24)
Second NP	651ms (24)	699ms (34)	675ms (20)	820ms (50)
Second Noun	511ms (15)	520ms (16)	524ms (12)	563ms (22)

Figure 3.16 depicts the outputs of the LME analyses for the articulation times of the first noun phrase (Figure 3.16 A) and the first noun (Figure 3.16 B). Participants showed significantly slower articulation times of the first noun phrase and noun in the low codable & one or more fixations condition in comparison to the intercept (noun phrase: $t = 2.9$, $p < .01$, noun: $t = 3.9$, $p < .001$). The difference between the high codable & no fixation condition and the intercept was not significant (noun phrase: $t = 1.2$, n.s., noun: $t = 0.1$, n.s.). The interaction was significant for the noun phrase ($t = 2.4$, $p < .05$) but not for the noun ($t = 0.2$, n.s.).

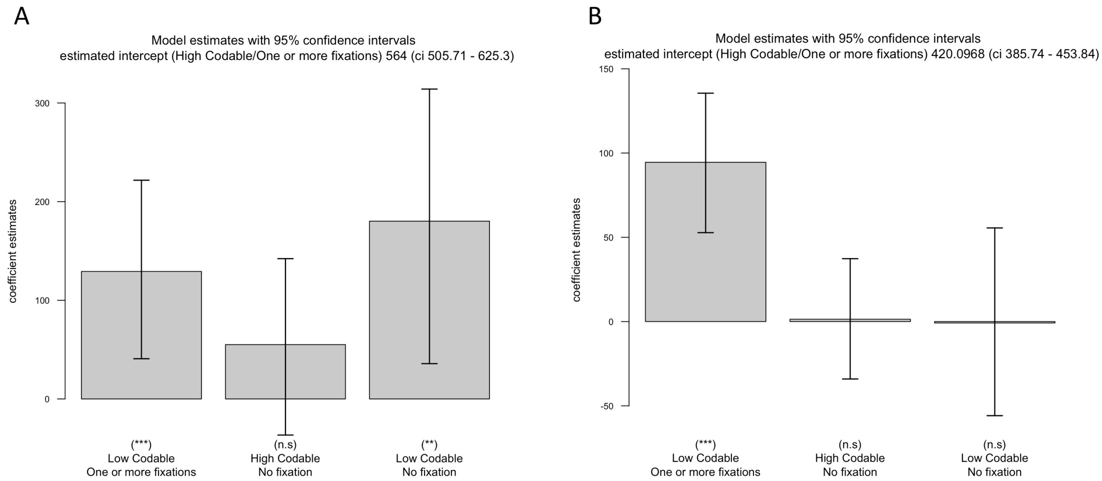

Figure 3.16 A: Articulation times of the first noun phrase. B: Articulation times of the first noun.

Pairwise comparisons showed a significant difference between high codable & no fixation and low codable & no fixation ($t = 3.8$, $p < .01$). Furthermore the difference between low codable & no fixation and low codable & one or more fixations was significant ($t = 3.4$, $p < .01$). The significant interaction for the first noun phrase indicates that the difference in articulation time between high and low codable items was higher for items in which the critical empty region was not fixated in comparison to items in which this region was fixated.

The graphs in Figure 3.17 illustrate the articulation times of the second noun phrase (Figure 3.17 A) and noun (Figure 3.17 B). No significant differences between low codable & one or more fixations and the intercept were found (noun phrase: $t = 1.0$, n.s., noun: $t = 0.3$, n.s.). Likewise, the difference between high codable & no fixation and the intercept was not significant (noun phrase: $t = 0.2$, n.s., noun: $t = 0.6$, n.s.). However, similar to the results for the first noun, the interaction was significant for the noun phrase ($t = 2.0$, $p < .05$) but not for the noun ($t = 1.4$, n.s.). The difference in articulation time between high and low codable items was higher for items in which the critical empty region was not fixated in comparison to items in which this region was fixated.

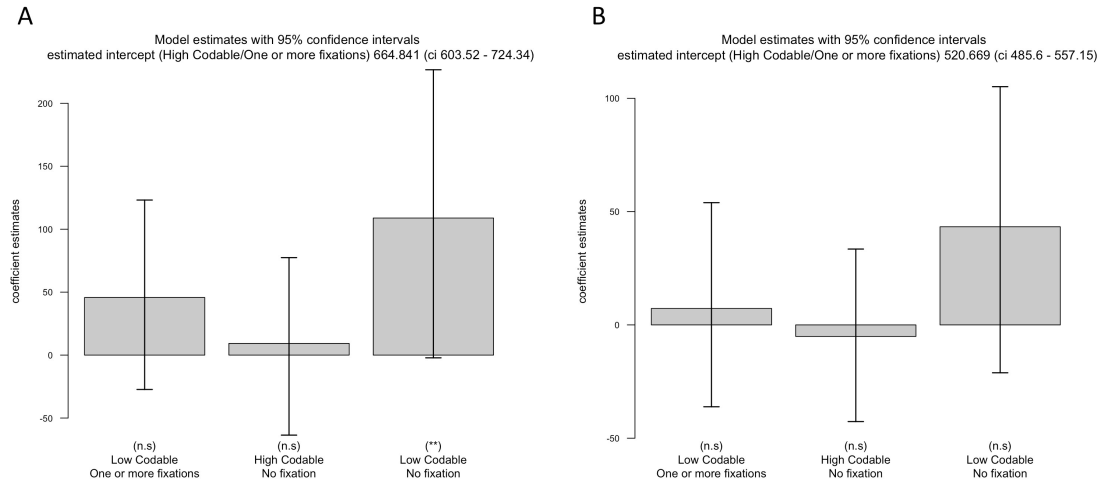

Figure 3.17 A: Articulation times of the second noun phrase. B: Articulation times of the second noun.

Pairwise comparisons showed a significant difference between high and low codable items in the no fixation condition (t = 3.0, p < .01). Furthermore, the difference between low codable & no fixation and low codable & one or more fixations was significant (t = 3.4, p < .01).

Speech onset times of the first and second noun phrases and nouns

The speech onset times of the first and second noun and noun phrase are depicted in Table 3.9.

Table 3.9: Speech onset times of the first NP and noun and of the second NP and noun for items in which the respective objects of the nouns were fixated. Values in parentheses represent standard deviations.

	One or more fixations		**No Fixations**	
Codability	**High**	**Low**	**High**	**Low**
First NP	1198ms (32)	1294ms (46)	1458ms (44)	1651ms (87)
First noun	1347ms (40)	1472ms (59)	1649ms (48)	2048ms (107)
Second NP	2419ms (60)	2884ms (107)	2783ms (81)	3102ms (146)
Second noun	2559ms (67)	3062ms (113)	2934ms (87)	3359ms (153)

Figure 3.18 Illustrates the speech onset times for the first noun phrase (Figure 3.18 A) and noun (Figure 3.18 B). The differences between high and low codability in items in which participants fixated the critical empty region was only marginally significant for the first noun phrase but significant for the first noun (noun phrase: t = 1.9, p < .10, noun: t = 2.1, p < .05).

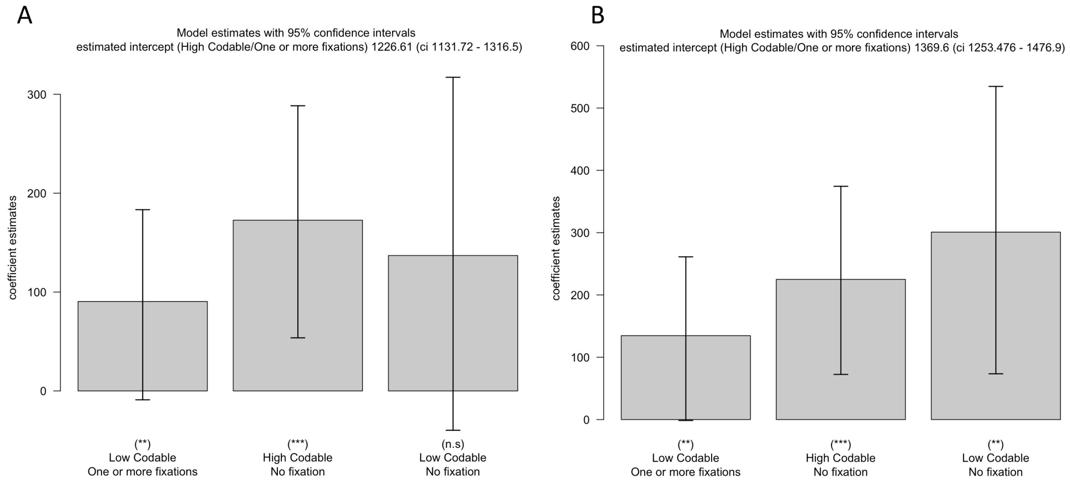

Figure 3.18 A: Speech onset times of the first noun phrase. B: Speech onset times of the first noun.

The difference between high and low codability in items with one or more fixations was significant in both the first noun phrase and the first noun (noun phrase: t = 2.8, p < .01, noun: t = 2.8, p < .01). The interaction was not significant for the first noun phrase but significant for the first noun (noun phrase: t = 1.6, n.s., noun: t = 2.5, p < .05). The significant interaction indicates that the difference in speech onset times between high and low codable items was higher for items in which the critical empty region was not fixated in comparison to items in which this region was fixated.

Finally, the speech onset times for the second noun phrase and noun are depicted in Figure 3.19 A and B, respectively. Participants showed significantly later articulation times of the second noun phrase and noun for the low codable items in comparison to the high codable items when they fixated the region in which an object had been (noun phrase: t = 4.3, p < .001, noun: t = 4.2, p < .001). Furthermore, in the high codability condition, speech onset times of the second noun phrase and the noun was significantly delayed when participants did not fixate the critical empty region in comparison to when they fixated it (noun phrase: t = 2.8, p < .01, noun: t = 2.7, p < .01).

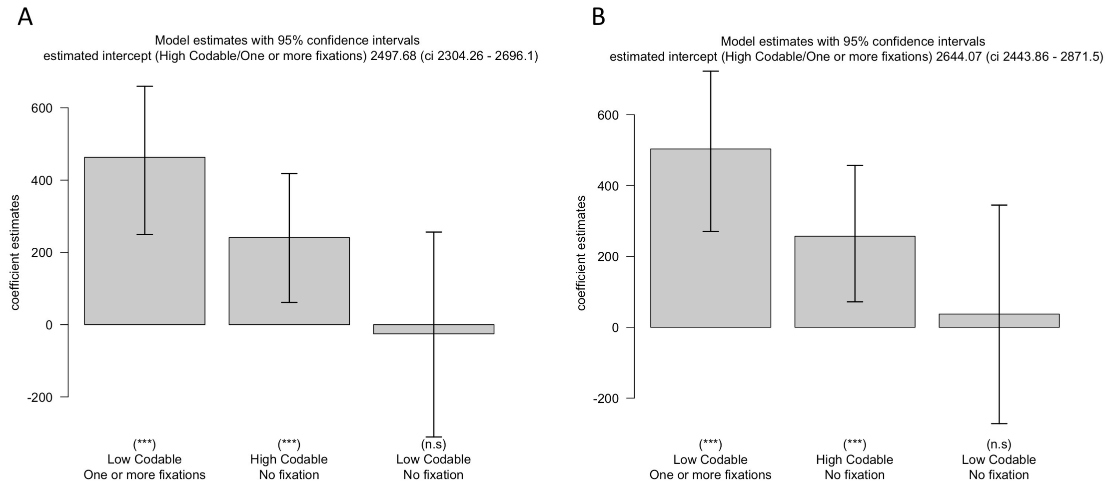

Figure 3.19 A: Speech onset times of the second noun phrase. B: Speech onset times of the second noun.

The interaction between Difficulty and Fixated was not significant for neither the second noun phrase nor the second noun (noun phrase: t = 0.2, n.s, noun: t = 0.3, n.s.).

3.4.3 Discussion

Experiment 2 tested two main questions. First, do speakers fixate empty regions previously occupied by an object when naming this object. Second, does fixating such an empty region facilitate language production?

The results of Experiment 2 showed that speakers fixated critical empty regions before naming objects. The number of fixations towards empty space was, however, not as high as when the objects were visible. Furthermore, fixations towards the second objects in the sentences were initiated earlier when the screen was empty in comparison to when it was not. Despite these differences, the results clearly show that name-related gazes are also initiated when the to-be named objects were not visible anymore. Most interesting, the effect between high and low codable objects, as investigated by Griffin (2001), could also be measured when the objects were not visible. Speakers were more likely to fixate an empty region when the associated objects were low codable in comparison to when the objects were high codable. The effect was however only found for the second but not for the first object. Codability was taken as a measure of difficulty for lemma retrieval. Thus, people were more likely to fixate an empty region when the lemma of an object, which had been located at this empty position, was harder to retrieve in comparison to an object with an easier to retrieve lemma.

In order to test whether fixating an empty critical region facilitated speech production, the articulation times and onset times of the critical objects were tested. Only items in which the objects did not appear again were tested. Items in which the critical empty regions were fixated were compared with items in which the critical empty regions were not fixated.

Concerning the articulation time, an interesting interaction between codability and fixated was found for the first and second noun phrase. The difference between high and low codable items was larger when speakers did not fixate the regions of the first and second objects in comparison to when these regions were fixated. Since speakers fixated low codable items more often than high codable items, these results suggest that fixating an empty region more often also provides more facilitation. In other words fixating the low codable objects more often than high codable objects equalises the level of difficulty in retrieval. However, when objects were not fixated, the differences in retrieval between high and low codable objects become more prominent.

Clear effects of speech onsets were found for the first and second nouns and noun phrases. Speech onset appeared later when speakers did not fixate the critical empty region in comparison to when they fixated it. A similar interaction effect as in articulation times was only found in speech onset times of the first noun.

These results clearly indicate that speakers fixate positions in which an object has been, when naming this objects. Difficult items are fixated more often than easier items. Evidence for a facilitation effect when fixating an empty critical region was found. Looking at an empty region results in both faster articulation and earlier speech onset times. Furthermore, spending more time on an empty critical region results in an additional facilitation effect.

3.5 General discussion of Experiments 1 and 2

In the first two experiments of this thesis, the interaction between visual and linguistic processing in language production was investigated. In previous experiments, it has been shown that speakers fixate to-be named objects for much longer than necessary in order to identify them (e.g. Meyer, et al., 1998). I proposed that these name-related gazes are a consequence of the close interplay between visual and linguistic information (Altmann & Kamide, 2007; Ferreira, et al., 2008). According to the model by Altmann & Kamide (2007), objects are stored as a multi-featured representation. When hearing a referent of an object all features are activated which leads to a shift of attention, and thus, a shift of gaze towards the object. I proposed that this mechanism also causes name-related gazes. Such a mechanism predicts that speakers fixate empty regions previously occupied by an object when naming this object. Furthermore, the model by Ferreira, et al. (2008) predicts that moving the eyes towards the location of an object or a region in which an object has been located causes a further activation of the object representation which leads to a facilitation of the linguistic feature.

Thus, two main predictions were made in the current language production study. First, speakers attend to an empty region previously occupied by an object when naming this object. Second, activation of the spatial location representation by either fixating the to-be named object in the correct location or by fixating an empty region previously occupied by this object leads to facilitation effects in language production.

The results from Experiments 1 and 2 confirmed these predictions. Experiment 2 showed that the looking at nothing effect, found in earlier studies in language comprehension (Altmann, 2004; Hoover & Richardson, 2008; Spivey & Geng, 2001), could also be found in language production. Speakers fixated regions in which to-be named objects have been located more often than alternative regions. Further results from Experiments 1 and 2 showed that fixating an object or a region previously occupied by an object facilitates articulation and speech onset times.

These findings support the assumption that name-related gazes are mediated by a model in which linguistic, visual, and spatial location features are tightly connected and that activating one of these features leads to a boost in activation of the whole multi-featured representation (Altmann & Kamide, 2007; Ferreira, et al., 2008). In more detail, in the apprehension phase of language production this multi-featured representation of an object is built. This procedure is very fast and can be achieved with only one fixation (Biederman, 1972; Biederman, et al., 1982; Potter, 1975). In the following formulation phase, the linguistic feature of the object is being activated which leads to a boost of the whole representation including the spatial location feature. The activated spatial location feature causes attention, and thus, the eyes to shift towards the position, or remain at the position in which the object has been. Fixating this position however, leads to an additional boost of the spatial location feature and thus, of the whole multi-featured representation including the linguistic feature. As a consequence, fixating the position of an object during language production results in a facilitation of language production (see also Bock, et al., 2003).

The spatial location feature seems to take a prominent role in guiding the eyes (Kahneman, et al., 1992). However, if compatible visual information is available, language production receives an additional boost in comparison to when incompatible visual information is perceived.

The study by Johansson, et al. (2006) found similar results such that speakers fixate empty regions during language production. In four experiments, participants were instructed to either retell a story or describe a scene. During retelling of the story or description of the scene, participants were either presented with an empty screen or sat in a completely darkened room. Johansson, et al. (2006) found that during language production people fixated locations in which objects have been or moved the eyes according to the spatial relation of the described objects or the objects within the scene. However, in contrast to the current experiments and to results from naming visible objects (Meyer, et al., 1998), participants usually fixated the locations in which objects have been after speech onset of the words. Thus, at the time when speakers fixated the relevant position, lemma and lexical retrieval of the language production process has already been completed.

What could be the reason for such different eye movement behaviour? In contrast to the experiments of the current study, Johansson, et al. (2006) presented the scenes or the stories for much longer than 500ms. Thus, when the scenes were presented, speakers were able to fixate objects for much longer than necessary to simply apprehend the objects. This should lead to a higher activated representation of the objects in comparison to the current study. I can only assume that this higher activation might trigger a different strategy. An already highly activated representation might not get a significant boost when the linguistic representation gets activated during production. Thus, a small boost might not attract eye movements as quickly as in the current experiments.

Other studies showed that the availability of visual information has no effect on the performance in speech production (Bock, et al., 2003; Griffin, 2004a; La Heij, van der Heijden, & Plooij, 2001). For example, Bock, et al. (2003) presented participants with clocks for either 100ms or 3000ms. This time difference did not cause a significant degradation of the ability of time telling. However, in the 100ms conditions, the authors did not compare items in which participants fixated the empty regions in which relevant information about the time has been located. Thus, it is difficult to tell which influences the looking at nothing effect had on this data. Furthermore, Bock, et al. (2003) acknowledge themselves that time telling is a different procedure than describing a scene. It is unlikely that time telling has been developed evolutionarily whereas judging the relation between objects might have been very important for survival. Furthermore, time telling is something humans do every day and is "well practised and highly idiomatic" (Bock, et al., 2003, p. 657). Thus, speakers might apply a different strategy in order to tell the time.

4 Chapter 4: The integration of visual and linguistic information in language comprehension

The current chapter investigates whether fixating regions previously occupied by an object facilitates processing of information related to this object. In contrast to the previous experiments, the effects were examined during language comprehension. The looking at nothing effect in language comprehension has been investigated in a number of studies (e.g. Altmann, 2004; Richardson & Spivey, 2000; Spivey & Geng, 2001; see also Section 2.1). Strong effects have been found showing that people fixate empty regions previously occupied by an object. In the current experiment, however, the focus was on whether fixating an empty region also facilitates language comprehension.

4.1 Overview of the experiments

The designs of the experiments of this study are very similar to Experiments 1 and 2 by Richardson & Spivey (2000). Participants heard four facts. These facts were termed the *introduction facts*, which were presented by an animated talking head. After hearing the introduction facts, participants heard a fact related to one of the introduction facts again. This fact was termed the *test fact* and the introduction fact related to this test fact was termed the *critical introduction fact*. The test fact was either a close repetition of the critical introduction fact or included a major change that caused the fact to be incorrect. Participants had to evaluate whether the test fact was correct or incorrect in relation to the critical introduction fact. However in comparison to Richardson & Spivey (2000), in Experiment 3 of the current study, participants did not see an empty screen while listening to the test fact. Rather, the same talking head that presented the introduction facts also presented the test fact and was located either at the same position in the two-by-two grid with respect to its position when it presented the introduction fact or at a different position. Assuming that participants fixate the talking head instead of looking at an empty portion of the grid (Krieger, Rentschler, Hauske, Schill, & Zetzsche, 2000;

Mackworth & Morandi, 1967), it was possible to compare the accuracy of evaluating the test facts when people fixated the same position with when they fixated a different position.

To preview the results of Experiment 3, I did not find differences in performance when the talking head was at the same in comparison to when it was at a different position. I hypothesised that this null effect might have been due to the possibility that the spatial location feature of the representation was updated by the visible talking head presenting the test fact. (For a more detailed explanation refer to the discussion of Experiment 3 and the general discussion.) Thus, in Experiment 4, I reverted to using the original empty screen paradigm by Richardson & Spivey (2000). However, I divided the facts into difficult and easy facts in order to investigate whether evaluating difficult facts draws more fixations to the associated empty port than easy facts. In Experiment 5, a mathematical problem was included between the four introduction facts and the test fact in order to increase the complexity of the task and to investigate the durability of the looking at nothing and facilitation effects.

4.2 Experiment 3: Looking at objects with varying locations during language comprehension

Experiment 3 tested whether retrieval of linguistic information related to an object is facilitated when people fixated the position of the object during encoding and retrieval in comparison to when they fixated a different position during retrieval. Similar to the experiments by Richardson & Spivey (2000), participants heard four introduction facts presented by a talking head and had to evaluate a subsequent test fact. In contrast to this study, however, I manipulated the location at which people were looking by presenting the talking head not only when the introduction facts were presented but also when the test fact was presented.

4.2.1 Methods

4.2.1.1 Participants

Twenty-one participants took part in Experiment 3. The average age of the participants was 22.8 (range 18-32). All participants were native speakers of British English and had normal or corrected to normal vision. Each participant was paid four pounds.

4.2.1.2 Material

Thirty-two items were generated. Each item consisted of four introduction facts. As in the study by Richardson & Spivey (2000) half of the items were world knowledge facts (Example (4.1)) and the other half were made up facts about fictional characters (Example (4.2)). A lab assistant collected these facts from Internet sources like Wikipedia.

(4.1) **World knowledge fact:** *The Bhut Jolokia, or Ghost Chilli, is the hottest chilli pepper in the world.*

(4.2) **Made up fact:** *Toby trapped a spider in his house and let it out through the window.*

A test fact was created from one of the four introduction facts. This test fact was either a close repetition of the critical introduction fact and was termed a *correct test fact*, or included a major change and was thus termed an *incorrect test fact*. Fifty percent of the test facts were correct and the other fifty percent were incorrect. Sentence (4.3) is an example of an incorrect test fact for Example (4.1).

(4.3) *The Red Savina habanero is the hottest chilli pepper in the world.*

Fifty percent of the test facts which tested the world knowledge facts were correct and the other fifty percent incorrect. The same allocation was applied to the test facts about fictional characters. All sentences were presented auditorily. The recording of the sentence was done in a sound recording studio. The files were saved as stereo wav files with a sample rate of 44100 kHz. The facts were presented by a talking head which was taken from the South Park TV series (character of Ike). The head was animated such that the mouth moved when the sound files were played. A talking head exhibited a size of 220 x 297 pixels on a set screen resolution of 1024 x 768 pixels. See Appendix C for a complete list of introduction and test facts.

4.2.1.3 Apparatus

An SR EyeLink1000 eye tracker with a sampling rate of 1000 Hz and a spatial resolution of less than 1/4 degree was used. The tracker was tower mounted and the eye movements of the right eye were tested. The stimuli were presented at a 19-inch CRT monitor running at 140Hz. The Experiment Builder software developed by SR research was used to run the experiment. The analysis was done using DataViewer, developed by SR Research Ltd. and Matlab.

4.2.1.4 Procedure

The experiment took place in a soundproof chamber. After reading the instructions, participants were seated in front of the monitor at a distance of approximately 90cm. Before the experiment started, participants were manually calibrated using a nine-point fixation stimulus. The EyeLink software validated the calibration. If validation was poor, it was repeated. The calibration procedure was repeated approximately two times during the experiment. Between items, participants had to fixate a point in the middle of the screen. This allowed the system to perform a drift correction if necessary.

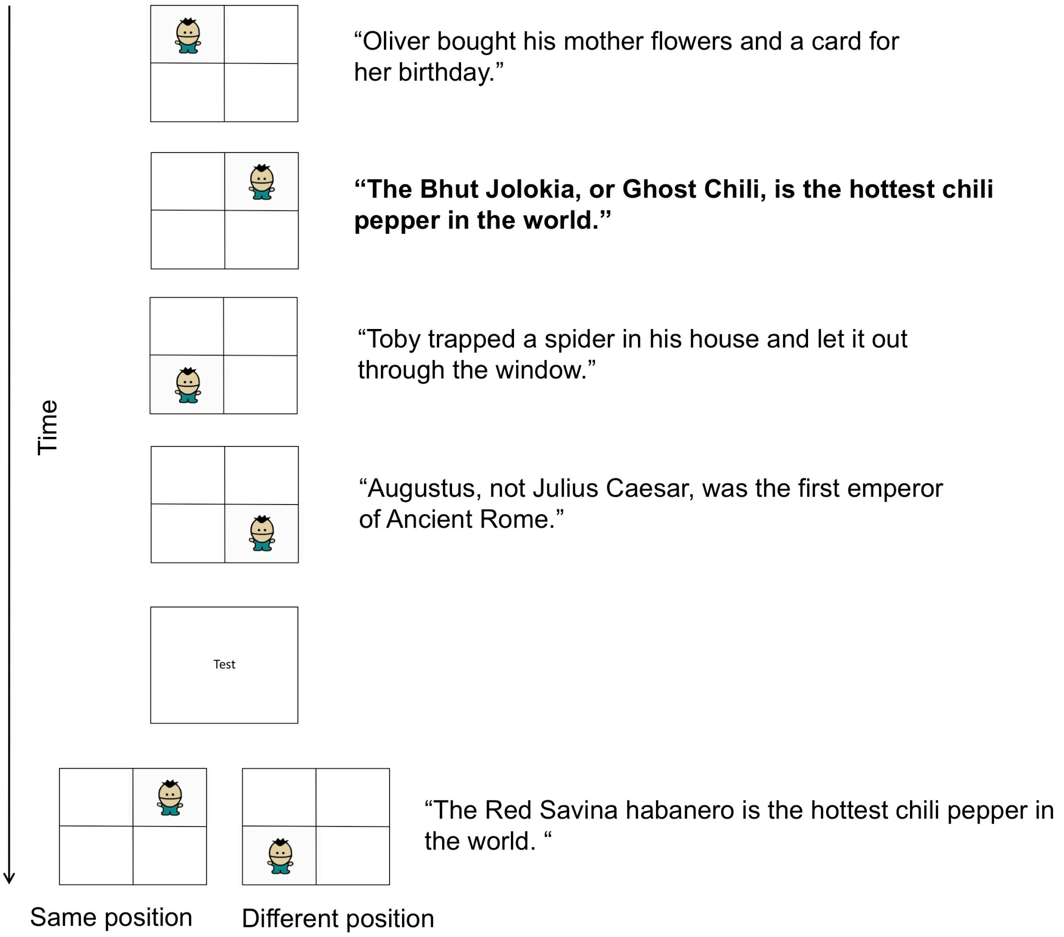

Figure 4.1: Schematic illustration of a trial in Experiment 3. The sentence in bold font denotes the critical introduction fact. The test fact (the last two grids) is an example for an incorrect test fact. Both conditions (same position, different position) are illustrated.

After participants fixated the fixation point in the middle of the screen, they were presented with a two-by-two grid consisting of equally sized four ports. The midlines of the grid were visible and formed a cross which filled the whole screen. The talking heads appeared in each of the ports in turn. The head was visible for 7500ms for each introduction fact. This time was divided in three sections. First, a 500ms time frame in which the mouth of the head did not move. Second, the time in which an introduction fact was presented. The head moved its mouth during this time. Third, the time between the offset of a fact and 7500ms in which the mouth did not move. After the head disappeared in one port, it immediately reappeared in a different port of the grid.

The sound files were presented over loudspeakers. After presentation of the four introduction facts the word TEST was presented in the middle of the screen to indicate the presentation of a test fact. Participants had to fixate this word in order to trigger the test fact. Thus, at the onset of the test fact, the eyes were always fixated on the screen centre. A talking head presented the test fact and was either at the same port as the critical introduction fact or at a different one. After the presentation of the test fact, participants evaluated it by pressing a left (correct) or right (incorrect) key on a button box. Before the experiment, participants were instructed to memorize the allocation of the keys and they were asked to hold the button box on their knees under a table in order to not look at the keys when answering the test fact. The order in which the talking heads appeared in different ports during the presentation of both the introduction and the test facts was counterbalanced. The experiment took approximately 30 minutes.

4.2.1.5 Data Analysis

To analyse the locations at which people fixated while evaluating the test facts the analysis of Richardson & Spivey (2000) was adopted. All fixations shorter than 100ms were deleted. In contrast to the study by Richardson & Spivey (2000), only fixations that landed on the exact area which the talking head covered were accepted as fixations in the corresponding port. In the remainder of this chapter, the term *port* denotes only the area covered by a talking head. Furthermore, I separately analysed the region around the fixation cross (centre of screen). This region was exactly in the middle of the screen and had a size of 100 pixels from left to right and 100 pixels from top to bottom. Fixations were measured from the speech onset of the test fact to the time the button was pressed to evaluate the test fact. In order to code the ports, the port in which the head presented the test fact was termed the talking head port. The remaining ports were coded Port 2, Port 3 and Port 4 going in a clockwise direction starting from the talking head port.

4.2.2 Results

4.2.2.1 Do people fixate the talking head?

When answering the test facts, participants almost always looked at the talking head. Figure 4.2 illustrates the mean number of fixations per trial on the different ports.

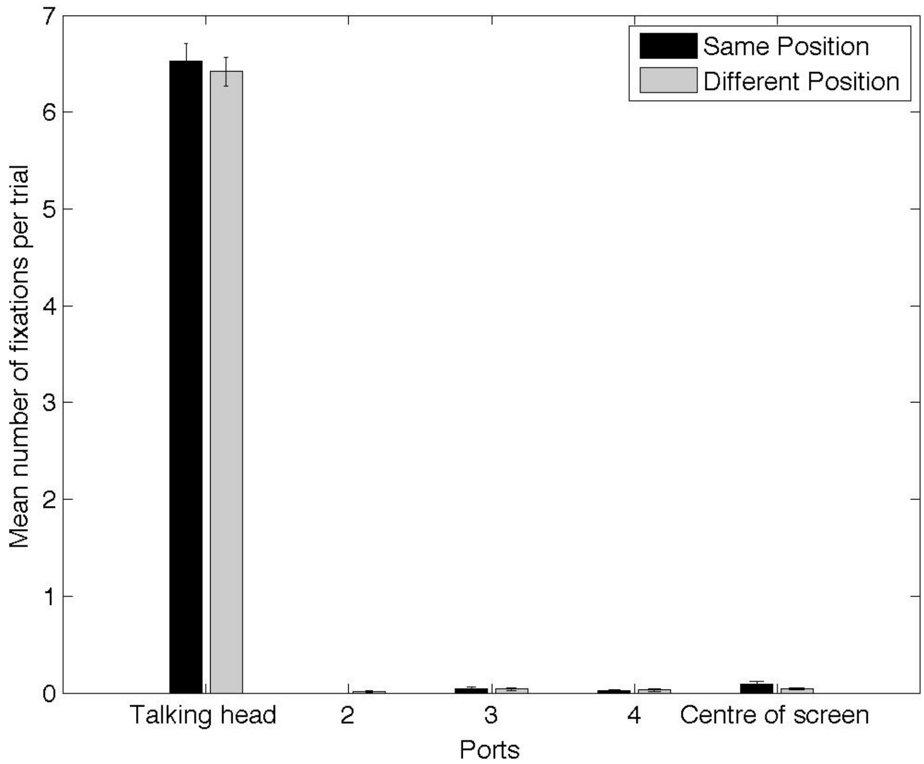

Figure 4.2: Experiment 3; Mean number of fixations per trial. The error bars denote standard errors.

To analyse the data, Ports 2 to 4 were combined into one region. Thus, a 3 (Fixation location: Figure Port, Port 2-4, Centre of screen) by 2 (Head location: Same Position, Different Position) design was utilised. A repeated measures ANOVA revealed a very strong and significant main effect of fixation location ($F_1(2,19) = 253.9$, $p < .001$, $F_2(2,30) = 492.5$, $p < .001$). The pairwise comparison analysis showed that the number of fixations in the port in which the talking head presented the test fact was significantly greater than fixations in all other ports and fixations in the centre of the screen (all ps <. 001; Talking head port: 6.47, Ports 2-4: 0.02, Centre of Screen: 0.07). The main effect of head location was not significant ($F < 1$; Same position: 1.34, Different position: 1.31). The fixation location x head location interaction was also not significant ($F < 1$).

4.2.2.2 Do people benefit from looking at the same location

To evaluate the performance of participants, two dependent variables were tested; percentage of correct responses and response time. Response time was measured from the onset of the test fact until the button press. Responses and response time of trials in which the talking head presented the introduction fact and the test fact in the same port were compared with trials on which the talking heads were in different ports. When the talking head was in the same port, 81.2% of the test facts were evaluated correctly versus 83.0% when the talking heads were in different ports. A t-test revealed no significant difference ($t = 0.6$). Similar results were found when the response time was compared. The difference between when the talking head was presented in same port (5720ms) versus in different ports (5545ms) was not significant ($t = 1.5$).

4.2.3 Discussion

While evaluating a test fact, participants mainly fixated the port in which the talking head presented the test fact, independent of whether the head was located in the port in which it presented the critical introduction fact. Almost none of the empty ports were fixated. No differences were found in either number of fixations at a port, the percentage of correct responses or the response time of evaluating the test fact between the same position and the different position conditions. These results appear not to be in agreement with the theory proposed by Ferreira, et al. (2008). According to this theory, participants should benefit from looking at the same port in which the introduction fact was presented. By looking at the same position in which the critical introduction fact was presented, the activated spatial location representation of the talking head should coactivate all other representations, including the linguistic representation, which should facilitate retrieval of the critical introduction fact.

However, since the talking head was visible in both conditions, the visual representation of the talking head, independent of its position, could also have coactivated the other representations. Contrary to Kahneman, et al. (1992), it has been shown that in addition to spatial location features, visual features are also able to access an objects file (Henderson, 1994; Henderson & Anes, 1994; Pollatsek, Rayner, & Henderson, 1990). The cause of the facilitation effect, predicted by Ferreira, et al. (2008) was not restricted to a reactivated spatial location feature. Therefore, a facilitation effect caused by the reactivated visual representation in both conditions might have concealed an effect caused by a reactivated spatial location representation in the same position condition. Furthermore, the spatial location information of the talking head representation might have been updated towards a new location (Kahneman, et al., 1992). In this case, no effect between the two conditions would have been expected because the head would have been located according to the location feature of the head's object file.

In order to circumvent this problem and to investigate the role of the spatial location representation, in Experiment 4, I reverted to using the original empty screen paradigm introduced by Richardson & Spivey (2000). However, in order to investigate whether the null effect in accuracy between looking at the critical location and not looking at the critical location found by Richardson & Spivey is due to a trade-off effect between fixation and difficulty, I also manipulated difficulty of the test facts.

4.3 Experiment 4: Looking at nothing during language comprehension

In Experiment 4 I tested whether the performance of evaluating the test facts improves when people look at the position in which the critical introduction fact was previously presented (the critical port). In contrast to Experiment 3, the talking heads were not present during the announcement of the test facts. Instead, participants looked at an empty screen during auditory presentation of the facts. The facts were either easy or difficult. More looks to the critical port were expected when the facts were difficult. Furthermore, similar to Richardson & Spivey (2000), participants should fixate the critical port more often than the other ports. According to Ferreira, et al. (2008), the performance of evaluating the facts should be better in items in which people fixated the critical port, in comparison to when people did not look at the critical port.

4.4 Method

4.4.1 Participants

Twenty-four participants took part in Experiment 4. Two participants were excluded due to problems with the calibration. The average age of the participants was 21.4 (range 18-29). All participants were native speakers of British English and had normal or corrected to normal vision. Each participant was paid four pounds.

4.4.1.1 Material

Thirty-two items each consisting of four introduction facts and one test fact were composed. In contrast to Experiment 3, the facts consisted only of world knowledge facts. Half of the facts were easy to answer because they were generally well-known facts (Example (4.4)) and the second half of the facts were not generally known to British undergraduates and thus more difficult (Example (4.5)).

(4.4) **Easy:** *Popeye is the comic strip character who eats spinach to increase his strength.*

(4.5) **Difficult:** *The Marabou stork, located in Africa, has the largest wingspan of any land bird.*

The level of difficulty was tested in a pretest. The procedure of the pretest was very similar to the procedure of Experiment 3. In each trial, participants heard four introduction facts followed by a test fact which tested one of the four introduction facts and was either correct or incorrect. However, neither of the introduction or test facts was presented by a talking head. Instead, participants looked at an empty screen. After they heard a test fact, they first had to evaluate it by pressing a left (correct) or right (incorrect) button on a button box. Following the evaluation, participants were asked to judge whether they heard the test facts before, or whether they were completely unfamiliar with it. Thirteen participants were tested. Sixteen items with the highest score in familiarity were allocated to the easy items and the sixteen items with the lowest score were allocated to the difficult items. On average, 87.0% of participants rated the easy items as familiar and only 16.3% rated the difficult items as familiar. Furthermore, correct responses in evaluating the facts were higher in the easy items (93.8%) than in the difficult items (78.4%; $t = 4.97$, $p < .001$). The response time for evaluating the test fact was faster for easy items (4806ms) than for difficult items (5134ms), but reached only marginal significance ($t = 1.91$, $p = .066$).

The critical word that distinguished a correct from an incorrect test fact was in 50% of the items at the beginning of the sentence and in 50% at the end of a sentence.

4.4.1.2 Apparatus

The same apparatus as in Experiment 3 was used.

4.4.1.3 Procedure

The procedure for Experiment 4 was almost identical to the procedure for Experiment 3. The only difference was that no talking head presented the test facts (see Figure 4.3). Apart from the two-by-two grid, the screen was completely blank.

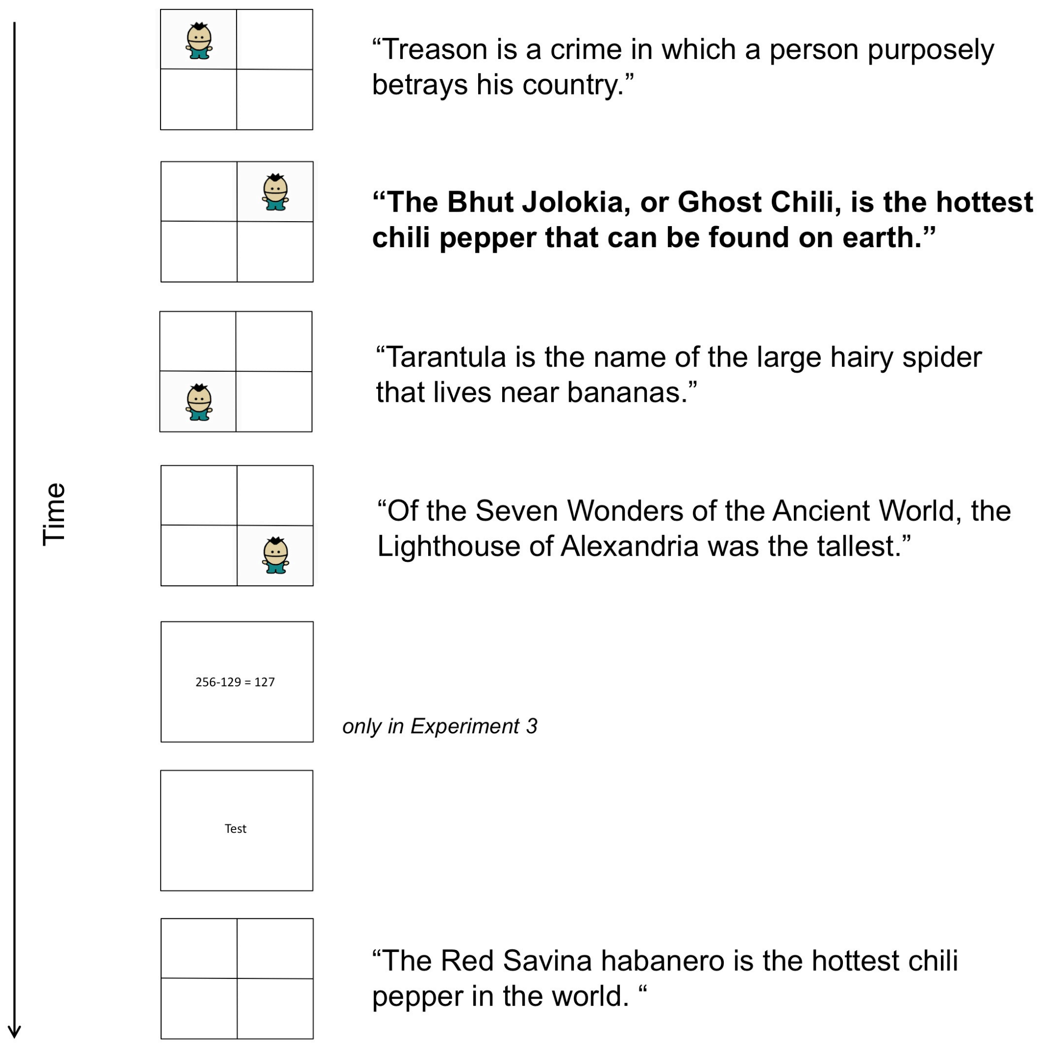

Figure 4.3: Schematic illustration of a trial in Experiment 4 and 5. The sentence in bold font denotes the critical introduction fact. The test fact (the last grid) is an example for a difficult incorrect test fact.

4.4.1.4 Data Analysis

Data analysis was similar to Experiment 3. However, in order to analyse differences between items in which participants fixated the port of the critical introduction fact while evaluating the test fact with items in which they did not fixate the same port, a linear mixed model was used (LME; Baayen, 2008). In order to analyse the number of fixations within the different ports, a similar coding scheme as in Experiment 3 was applied. However, instead of using the head port as the reference port, the port in which the critical introduction fact was introduced was the reference port. The remaining ports were coded Port 2, Port 3 and Port 4 going in a clockwise direction starting from the port of the critical introduction fact. As in Experiment 3, fixations in regions in which the fixation cross has been (centre of screen) were analysed as well. Easy and difficult items were also compared with each other.

4.4.2 Results

4.4.2.1 Do people look at nothing

Figure 4.4 illustrates the mean number of fixations per trial on the four ports and on the centre of the screen.

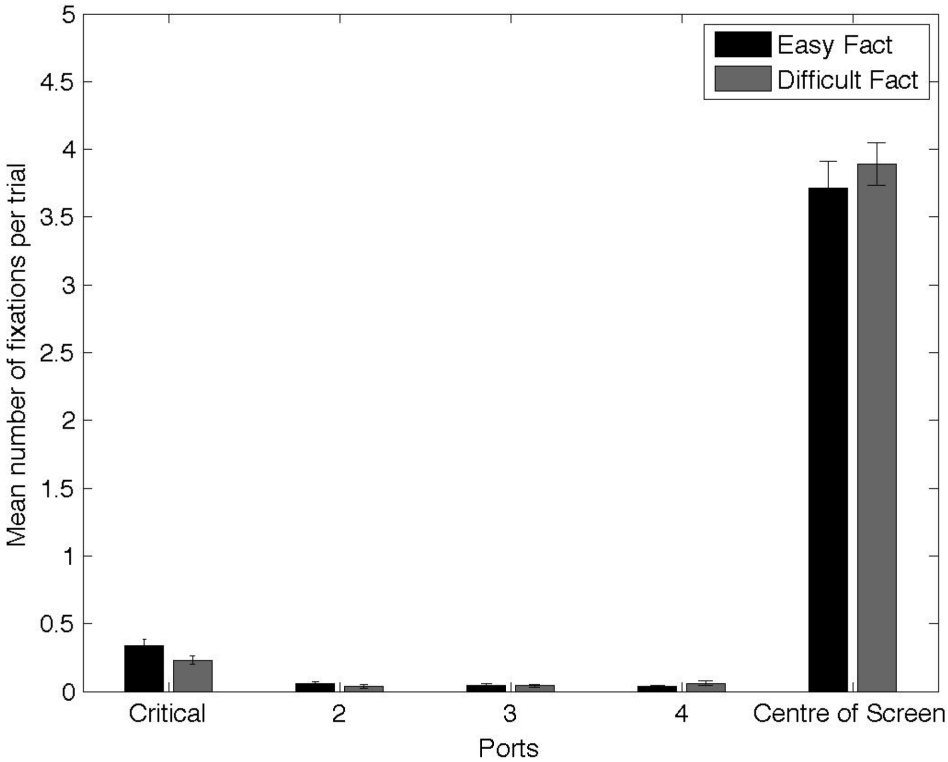

Figure 4.4: Experiment 4; Mean number of fixations per trial. The error bars denote standard errors.

To analyse the data, Ports 2 to 4 were combined into one region. Thus, a 3 (Fixation location: Critical Port, Port 2-4, Centre of Screen) by 2 (Difficulty:

Easy, Difficult) design was utilised. A repeated measures ANOVA revealed a significant main effect of fixation location ($F_1(2,20) = 76.2$, $p < .001$, $F_2(2,29) = 347.0$, $p < .001$). Pairwise comparisons showed that people fixated more often at the critical port than at Port 2-4 ($p < .05$; Critical port: 0.28, Ports 2-4: 0.05). Furthermore, participants fixated more often at the centre of the screen than at the critical port or at Port 2-4 (ps < .001; Centre of the screen: 3.80). The main effect of difficulty (Easy fact: 0.84, Difficult Fact: 0.85) and the fixation location x difficulty interaction were not significant (Fs<1.9).

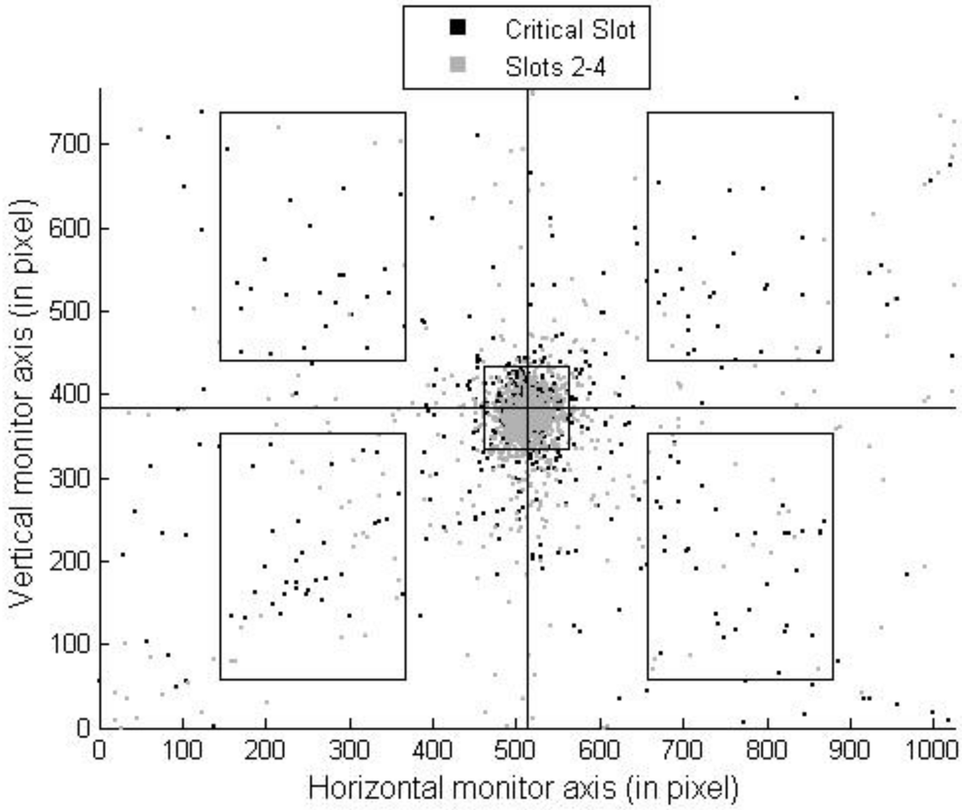

Figure 4.5: Distribution of fixations during the evaluation of the test facts. The squares illustrate the positions in which the talking heads were located. Since the location of the critical port was randomized, fixations towards a critical port are shown by dark marks and fixations towards ports 2-4 by brighter marks.

In order to illustrate the distribution of fixations on the whole computer screen during evaluation of the test facts refer to Figure 4.5. This scatter plot illustrates that fixations were strongly clustered around the centre of the monitor. Furthermore, the exact regions in which the talking heads were located did not receive additional attention. Thus, the looking at nothing effect was very weak in this experiment.

4.4.2.2 Do people benefit from looking at the same location

As in Experiment 3, performance was tested by measuring the proportion of correct responses and the response time. Response time was measured from the onset of the test fact until the button of the button box was pressed.

Table 4.1: Experiment 4; correct responses and response times of the evaluation of the test facts.

Condition	Correct Responses	Response Time
No Fixation & Easy	94.6%	5072ms
No Fixation & Difficult	83.1%	5169ms
One or more fixations & Easy	91.4%	4850ms
One or more fixations & Difficult	82.4%	5072ms

Table 4.1 shows the proportion of correct responses and the response time in all four conditions. A linear mixed model (Jaeger, 2008) was used to analyse the responses and the response time. Such a model was chosen because I compared items in which people fixated the critical port with at least one fixation (One or more fixations) with items in which they did not fixate the critical port (No fixation). The number of 'no fixation' items and 'one or more fixations' items were not evenly distributed and a mixed effects model can handle such unbalanced datasets. Two fixed effects were used: Fixated (No fixation, One or more fixation) and Difficulty (Easy, Difficult). The condition no fixation & easy comprised the intercept.

Figure 4.6 illustrates the outcome of the mixed logit model analysis for the responses.

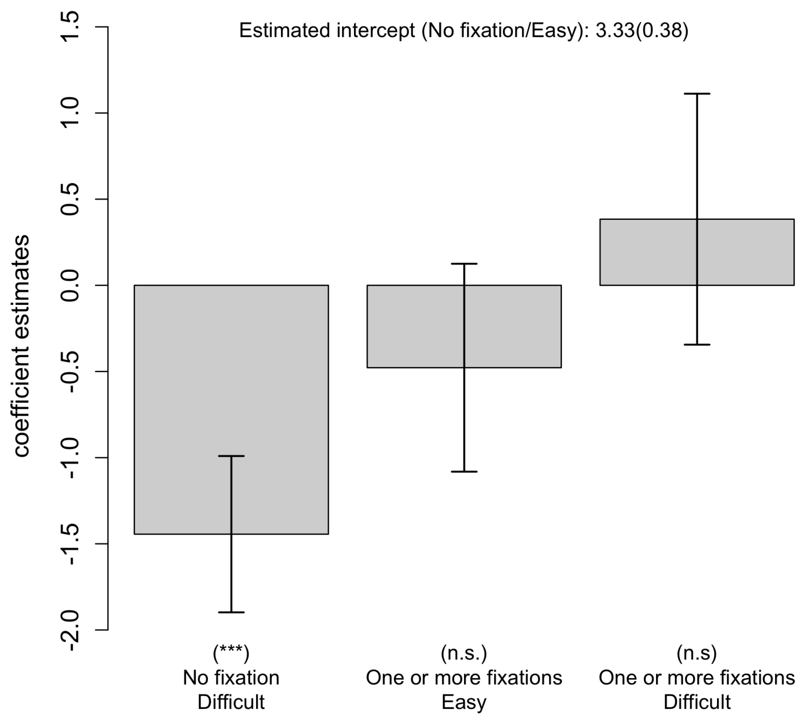

Figure 4.6: Experiment 4; linear mixed logit model of the correct responses of the evaluation of the test facts. The error bars denote estimated standard errors.

Participants produced correct responses less often when presented with difficult items in comparison to easy items (z=3.18, p < .01). No difference in numbers of correct responses between items in which participants fixated the critical port in comparison to when they did not fixate the critical port was found (z=0.79, p = .42). Furthermore, no significant interaction was found (z=.53, p = .60).

To analyse the response time I applied a linear mixed model for continuous output variables. Only items that were correctly evaluated were included in this analysis. Figure 4.7 illustrates the output of the model.

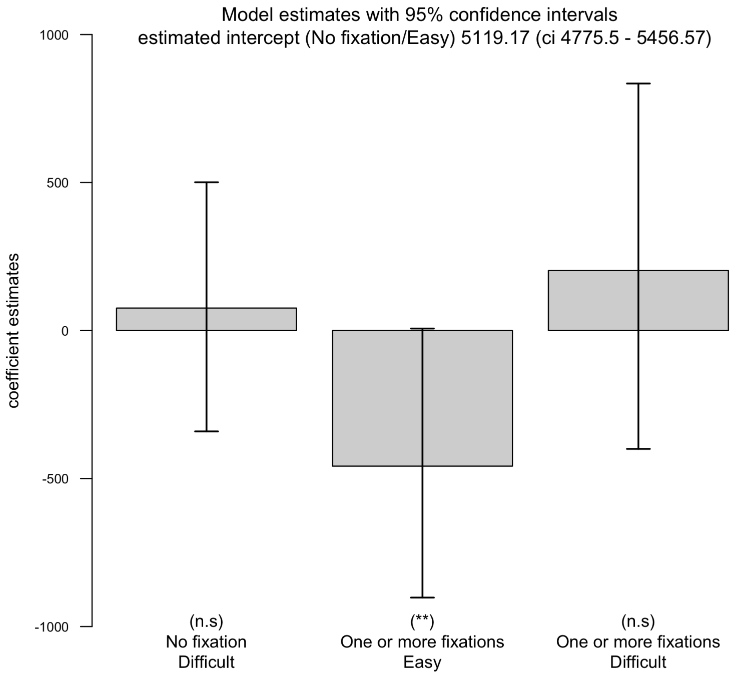

Figure 4.7: Experiment 4; linear mixed model of response times of the evaluation of the test facts. The error bars denote 95% confidence intervals (derived from a Marcov Chain Monte Carlo simulation).

Participants did not differ in response time when they evaluated difficult test facts in comparison to easy test facts ($t = .33$, $p = .74$). However, a result important for our hypothesis is that participants were faster when they looked at the critical port in comparison to when they did not ($t = 2.1$, $p < .05$). The interaction between fixated x difficulty was not significant.

4.4.3 Discussion

In Experiment 4, results from Richardson & Spivey (2000) were replicated. Participants were more likely to look at the critical port during evaluation of the test facts. Furthermore, no differences in response accuracy were found between items in which participants fixated the critical port in comparison to items in which they did not fixate the critical port. This null effect was not due to a trade-off effect between fixation and difficulty. Differences in viewing behaviours between difficult and easy items were not found. Thus, items that are more difficult are not fixated more often which might have improved the response accuracy. However, in the current experiment the response time of evaluating the test facts was also analysed. Most interestingly, participants were faster when the critical port was fixated in comparison to when it was not fixated. Thus, Experiment 4 presents evidence that reactivating the spatial location representation influences processing of information related to an object previously located at this location.

The analysis of Experiment 4 differed in some ways from the analysis by Richardson & Spivey (2000). Only fixations that were located in the exact area in which the talking head appeared were included in the analysis. Furthermore, participants were asked to fixate the middle of the screen between the presentation of the four introduction facts and the test fact. This region was analysed separately. In comparison to the study by Richardson & Spivey, I found fewer fixations per trial within the ports. This dissimilarity was probably due to the fact that a smaller region was analysed in the current study. Most intriguing, however, was that participants had a very strong tendency to remain at the centre of the screen during the evaluation of the test facts. This indicates that although people benefit from looking at the critical port, they rather remain at a given location and not move their eyes at all.

A possible cause for the reported null effect in fact difficulty could be that the difference between easy and difficult facts was not large enough. The high accuracy for the easy items might suggests that a ceiling effect could be the reason for this low difference. Therefore, in Experiment 5, a mathematical problem was inserted between the presentation of the four introduction facts and the test fact. This problem had to be evaluated. I assumed that due to an increased demand of working memory evaluating the test fact is more difficult in Experiment 5. Furthermore, I was able to test the durability of the effect reported in Experiment 4.

4.5 Experiment 5: Looking at nothing after engaging in an additional task

In Experiment 5 the same hypotheses as in Experiment 4 were tested. To impede the evaluation of the test facts, a mathematical problem was inserted between the presentation of the introduction facts and the test fact. Furthermore, the grid, which was visible during presentation of the test fact in Experiment 4, was removed. In Experiment 4, most of the fixations remained at the centre of the screen. Since the grid formed a cross that was centred in the middle of the screen, this location was highly salient on an otherwise empty screen. Therefore, participants' fixations might have been drawn to this location.

4.5.1 Method

4.5.1.1 Participants

Twenty-four participants took part in Experiment 5. The average age of the participants was 21.3 (range 18-27). All participants were native speakers of British English and had normal or corrected to normal vision. Each participant was paid four pounds.

4.5.1.2 Material

The same sentences as in Experiment 4 were used. The mathematical problems were always subtractions (e.g. 366 – 119 = 247). The last digit of the minuend (366) was always smaller than the last digit of the subtrahend (199) and the equation was thus considerably difficult to solve. The numbers included in the problems were chosen randomly but consisted always either of two or three digits. For a list of the mathematical problems see the Appendix E.

4.5.1.3 Apparatus

The same apparatus as in Experiment 3 was used.

4.5.1.4 Procedure

Apart from two changes, the procedure of Experiment 5 was identical to the procedure of Experiment 4. First, after the four introduction facts were introduced, a mathematical problem was included (see Figure 4.3). The mathematical problem was presented in the centre of the screen. Participants had to evaluate the mathematical problem by either pressing the left key on the button box if the problem was correct or the right key if the problem was incorrect. The mathematical problem remained on screen for 8 second but disappeared earlier if the participants pressed the left or right button earlier. Second, the lines of the two-by-two grid that were visible in the test screen of Experiment 4 were omitted. This was done because in Experiment 4 fixations very often remained at the middle of the screen. A reason for this behaviour could be the saliency of this region due to the convergence of the two lines of the grid. Thus, in Experiment 5, participants were presented with an utterly empty screen when hearing the test facts.

4.5.1.5 Data Analysis

To analyse the number of fixations within the different ports, the same coding scheme as in Experiment 4 was applied. Fixations at the four ports and at the region at the centre of the screen were analysed.

4.5.2 Results

4.5.2.1 Do people look at nothing?

Figure 4.8 illustrates the mean number of fixations per trial on the four ports and at the centre of the screen.

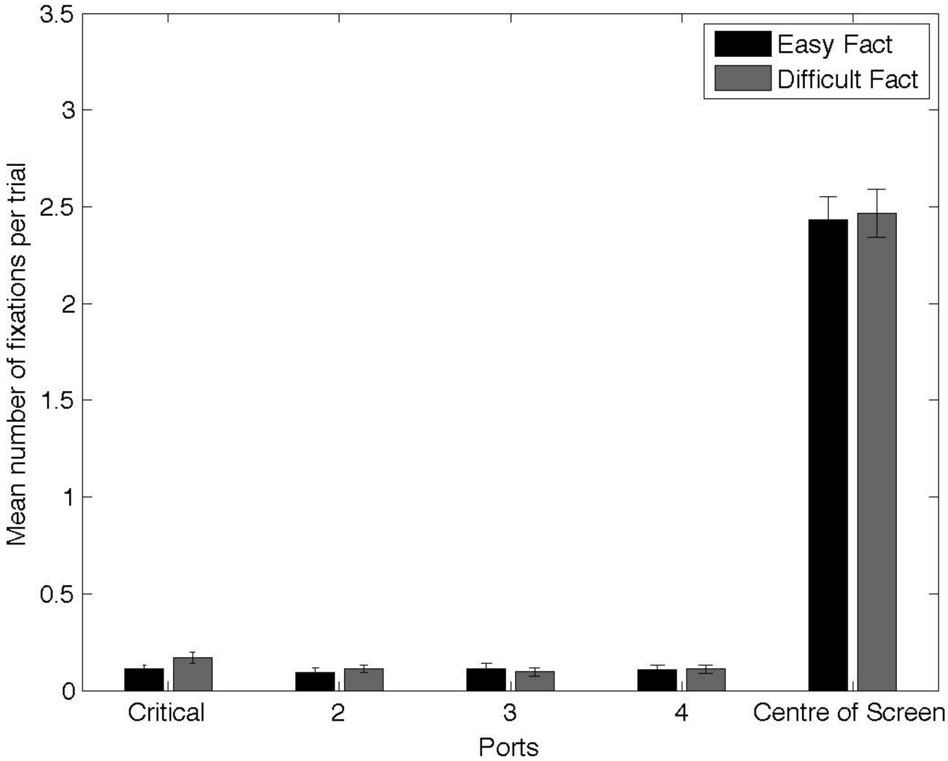

Figure 4.8: Experiment 5; Mean number of fixations per trial. The error bars denote standard errors.

To analyse the data, Port 2 to 4 were combined into one region. A 3 (Fixation location: Critical Port, Port 2-4, Centre of screen) by 2 (Difficulty: Easy, Difficult) design was utilised. A repeated measures ANOVA revealed a significant main effect of fixation location ($F_1(2,22) = 21.7$, $p < .001$, $F_2(2,29) = 462.2$, $p < .001$). Pairwise comparisons showed that people did not fixate more often at the critical port than at Ports 2-4 ($F < 1.2$; Critical port: 0.14, Ports 2-4: 0.11). However, participants fixated more often at the centre of the screen than at the critical port or at Ports 2-4 ($ps < .001$; Centre of the screen: 2.45). The main effect of difficulty (Easy fact: 0.56, Difficult Fact: 0.59) and the fixation location x difficulty interaction was not significant ($Fs < 1.1$).

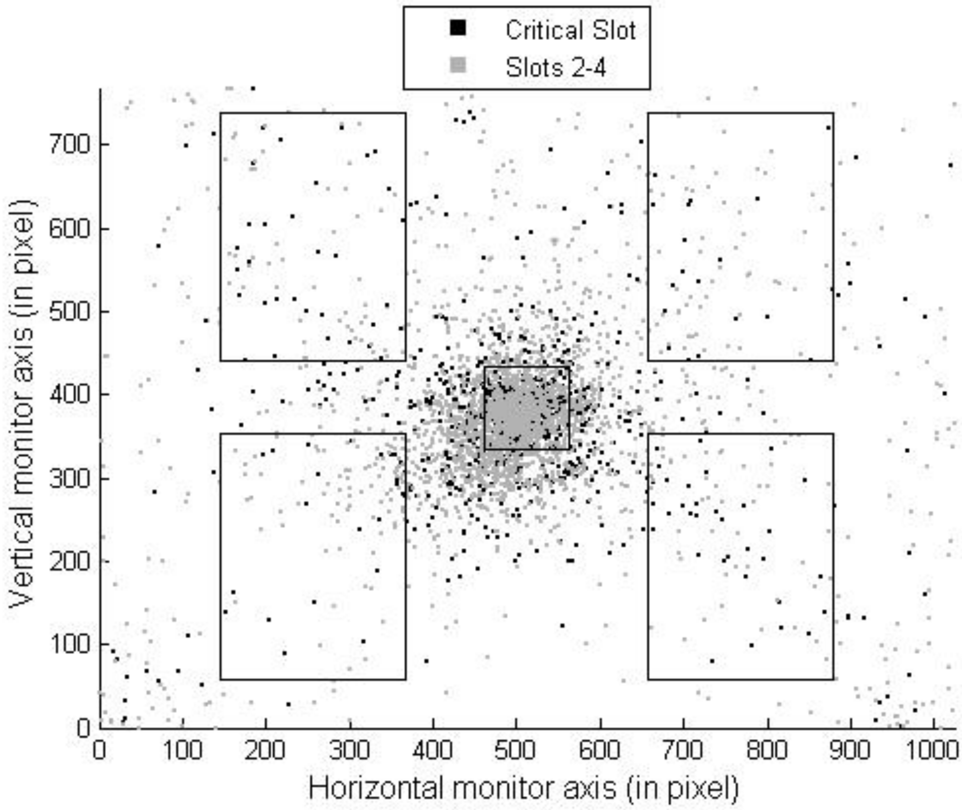

Figure 4.9: Distribution of fixations during the evaluation of the test facts. The squares illustrate the positions in which the talking heads were located. Since the location of the critical port was randomized, fixations towards a critical port are shown by dark marks and fixations towards ports 2-4 by brighter marks.

As in Experiment 4, the distribution of fixations on the whole computer screen during the evaluation of the test facts was plotted (see Figure 4.9). Fixations were also clustered around the centre of the screen. Furthermore, the exact regions in which the talking heads had been located did not receive additional attention. However, fixations were not as densely clustered around the centre of the screen as in Experiment 4. This observation is confirmed by the comparison of the numbers of fixations at the centre of the screen in Experiment 4 and 5 (Figure 4.4 and Figure 4.8). In Experiment 4, participants fixated the centre of the screen on average 3.73 times, and in Experiment 5 only 2.45 times.

4.5.2.2 Do people benefit from looking at the same location

As in Experiment 4, the proportion of correct responses and the response time were measured. Table 4.2 shows the proportion of correct responses and the response time in all four conditions.

Table 4.2: Experiment 5; correct responses and response times of the evaluation of the test facts.

Condition	Correct Responses	Response Time
No Fixation & Easy	95.4%	5086ms
No Fixation & Difficult	84.2%	5110ms
One or more fixations & Easy	93.8%	4836ms
One or more fixations & Difficult	67.4%	5339ms

The responses were analysed by using a mixed logit model and the response times were analysed with a general linear mixed effects model. Two fixed effects: Fixated (No fixation, One or more fixation) and Difficulty (Easy, Difficult) were used. Participants and items were included as a random factor. The condition no fixation & easy comprised the intercept. Figure 4.10 shows the output of the logit model of the responses.

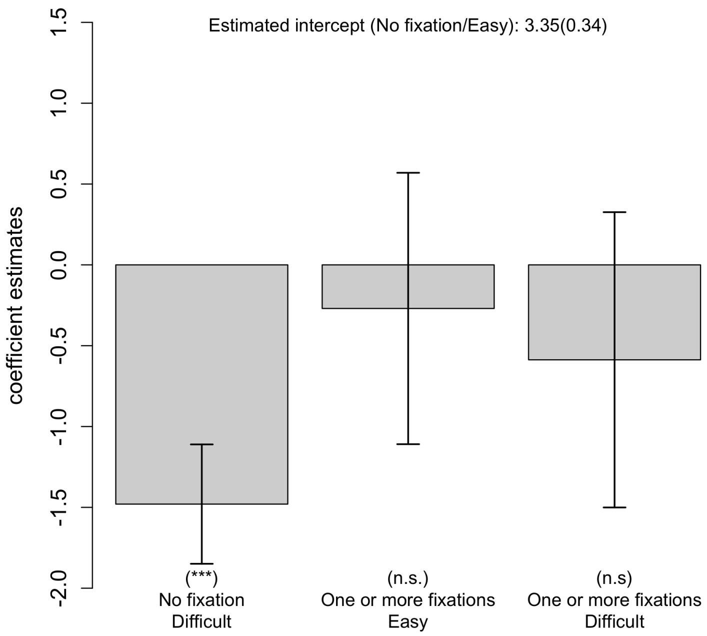

Figure 4.10: Experiment 5; linear mixed logit model of the correct responses of the evaluation of the test facts. The error bars denote estimated standard errors.

Participants showed significantly fewer correct responses when presented with difficult items in comparison to easy items ($z = 4.01$, $p < .001$). No difference in numbers of correct responses between items in which participants fixated the critical port in comparison to when they did not fixate the critical port was found ($z = 0.32$, $p = .75$). Furthermore, no significant interaction was found ($z = .64$, $p < .52$).

To analyse the response time of evaluating the test facts, only items that were correctly evaluated were included. Figure 4.11 illustrates the output of the linear mixed model.

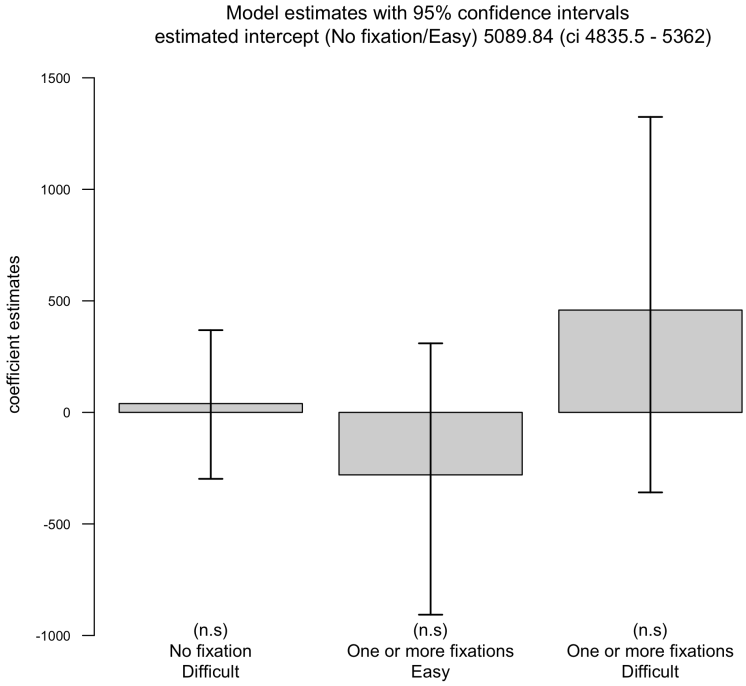

Figure 4.11: Experiment 5; Linear mixed model of response times of the evaluation of the test facts. The error bars denote 95% confidence intervals (derived from a Marcov Chain Monte Carlo simulation).

Participants did not differ in response time when the test facts tested a difficult fact in comparison to an easy fact (t = .25, p = .80). Furthermore, in contrast to Experiment 4, participants were not faster when they looked at the critical port in comparison to when they did not look (t = .90, p = .37). The interaction between fixated x difficulty was also not significant (t = 1.90, p = .37).

4.5.3 Discussion

In the current experiment, an additional task was included between the presentation of the introduction facts and the test fact. Participants had to additionally evaluate a mathematical problem. Furthermore, the grid was removed during presentation of the test facts. Thus, participants looked at a completely empty screen.

The mathematical problem was included in order to increase the complexity of the experiments. However, error rate was only 2.5 % higher than in Experiment 4. The additional mathematical problem, however, changed the results in a different way. Speakers were not more likely to fixate the critical ports more often than the other ports during evaluation of the test facts. Furthermore, participants did not show a faster response time when they fixated the critical port in comparison when they did not fixate it. I therefore propose that the effects reported by Richardson & Spivey (2000), and in Experiment 4 of the current study are relatively short-lived.

4.6 General discussion of Experiments 3, 4, and 5

The experiments in this chapter investigated the integration of visual and linguistic information in language comprehension. Two main questions were investigated. First, do people fixate locations, related to information, when retrieving this information from memory? Second, does reactivation of the spatial location representation of an object facilitate retrieval of information associated with this object?

In three experiments, participants heard four facts presented by a talking head appearing in the four different ports of a two-by-two grid (analogous to Richardson & Spivey, 2000). Thus, the linguistic information conveyed by the introduction facts were associated with a talking head in a specific port within the two-by-two grid. After the presentation of these introduction facts, participants heard a test fact which was either correct or incorrect in relation to the critical introduction fact. The correctness of this test fact had to be evaluated. In Experiment 3, the test facts were presented by talking heads located either in the same port as the critical introduction fact or in a different port. By showing the talking head in either the same or in a different port during the presentation of the test fact, the spatial location representation of the talking head associated with the critical introduction fact was either reactivated or not reactivated. In Experiments 4 and 5, the test facts were presented without a talking head. People only saw an empty two-by-two grid (Experiment 4) or a totally empty screen (Experiment 5). Thus, in Experiments 4 and 5, the spatial location representation of the talking head associated with the critical introduction fact was reactivated without the presentation of a visual trigger but by a non-triggered gaze shift of the participants. In addition, the level of difficulty of the introduction facts was manipulated in Experiments 4 and 5. Furthermore, in Experiment 5, a mathematical problem was inserted between the presentation of the introduction facts and the test fact in order to investigate the durability of the facilitation effect.

Of major interest in all three experiments was whether people fixate the location of the critical port (i.e. the port in which the critical introduction fact was presented) and whether they show an increased number of correct responses and faster response times when the location of the critical port was fixated.

4.6.1 Fixation of the location of the critical port

In Experiment 3, participants fixated the head that announced the critical fact. This result was not very surprising. People have a strong tendency to fixate objects on an otherwise empty screen (Krieger, et al., 2000; Mackworth & Morandi, 1967).

Of more interest were the results from Experiment 4 and 5. In Experiment 4, the results from Richardson & Spivey (2000) were replicated. Participants fixated the empty critical port while evaluating the test fact associated with this port more often than other ports. These results present evidence for the theory by Altmann & Kamide (2007). During retrieval of linguistic information, the linguistic feature of a multi-featured representation of an object is activated which coactivates the spatial location feature. This leads to a shift of attention and, thus, people fixate the relevant location even if it is empty.

However, this effect was relatively weak. The critical port received less than 0.4 fixations per trial. Thus, in many trials it was not fixated at all. Furthermore, most of the fixations were located at the centre of the screen. Since participants were asked to fixate the centre of the screen between the introduction facts and the test fact, they often did not move at all during presentation of the test facts. Similar results were also found by Spivey & Geng (2001). Furthermore, these results are not as different from the results by Richardson & Spivey (2000). In this study, the number of fixations within the critical port was never more than one and could be as low as 0.5. Furthermore, Richardson & Spivey (2000) measured fixations not at the exact regions of the heads but only at a whole quarters of the screen. This might account for the higher number of fixations. The distribution of fixations at the screen also showed that the regions in which the talking heads had been located did not receive additional attention. Fixations were evenly distributed over the whole screen and were only strongly clustered around the centre of the screen.

In addition, Experiment 5 showed that the looking at nothing effect seems to be very short-lived. If participants engage in a different task between encoding and retrieval of information, the empty critical region was not fixated more often than alternative regions.

Before interpreting these results within the model of integrated visual and linguistic information, the facilitation effects of looking at empty regions are summarised.

4.6.2 Do people benefit from looking at critical regions

The results of Experiment 3 demonstrated that retrieval of linguistic information is not facilitated by the reactivation of the spatial location feature. The number of the correct responses and the response times did not differ between conditions in which the talking head presented the test facts in the critical port in comparison to when it presented the test fact in a different port. However, since the talking head was visible, not only the spatial location representation of the head was activated but also the visual representation. Even though the spatial location feature plays a major role in accessing objects' features (Kahneman, et al., 1992), it has been shown that other representations are also involved (Henderson, 1994; Henderson & Anes, 1994; Pollatsek, et al., 1990). In an experiment similar to Kahneman, et al. (1992), Henderson (1994) presented participants with two letters appearing in a preview screen. In a following target screen the letters changed their positions, their form, or the flanked letter. Participants had to name the letter in the target screen that also appeared in the preview screen. Different from the study by Kahneman, et al. (1992), people had to execute a saccade toward the letters. Henderson (1994) found a facilitation effect even if the letter changed the position. Thus, the visual features of an object independent of its position caused a facilitation effect. It is therefore possible that the visual features of the talking heads, appearing at a different position, also facilitated language processing and might have distorted a facilitation effect originating from the spatial location representation. Furthermore, Kahneman, et al. (1992) showed that the spatial location features are updated when objects are moving. A head appearing at a non-critical port might have been perceived as a movement from the critical port and thus received an update of the spatial location. Under these circumstances, a talking head appearing at the critical port would show no facilitation effect in comparison to a head appearing at a different port.

Therefore, in Experiment 4 and 5, the test facts were presented without a talking head. As mentioned in Section 2.1, Richardson & Spivey (2000) did not find any facilitation effects when participants fixated the critical port in comparison to when they did not fixate it during fact evaluation. However, a possible trade-off effect between fixation and difficulty might have caused this null-effect. Therefore, in the current study the level of difficulty of the facts was varied. I did not find differences in number of fixations on the critical port between easy and difficult items. Furthermore, in Experiment 4, I replicated the null-effect of accuracy between items in which the critical port was fixated in comparison to when it was not fixated, as found by Richardson & Spivey (2000). However, a significant effect of response time was found. People were faster in evaluating the test facts when they fixated the empty critical port in comparison to when they did not fixate it. Richardson & Spivey (2000) did not find any facilitation effects. However, the response time was not reported. In Experiment 5, no facilitation effects were found.

Taken as a whole, the results of the current study presented evidence that retrieval speed of linguistic information is facilitated when the spatial location of an object, associated with this linguistic information, is reactivated. This effect is however relatively short-lived and relies on the two occurrences of the object being in close temporal proximity without any intervening task.

4.6.3 Revisiting the looking at nothing effect in language comprehension

To summarise, evidence for a model with an integrated linguistic and visual information in memory was found. Participants fixated empty regions that were previously occupied by the talking heads when processing information conveyed by the talking heads. Faster response times were measured when people fixated the empty region in comparison to when they did not.

Thus a bound representation including visual, spatial, and linguistic information was created. When the information of the fact was retrieved, the spatial location feature of a head was activated which guided the eyes to move to the location of the head. By fixating this region, the spatial location feature received an additional boost which coactivated the whole representation and facilitated retrieval of the facts.

However, a number of refinements have to be made to the model. First of all, the link between linguistic representations and the spatial location representations within the model does not seem to be very stable and pronounced. The results from Experiments 4 and 5 showed that the critical empty port was fixated only in about a third of the trials. Thus, activation of the linguistic information does not cause a strong boost of the spatial location representation. Furthermore, the effect is not long lived. The effect disappears if people engage in a different task between encoding and retrieval of information. As a consequence of the relatively weak link between the linguistic and the spatial location representation, the facilitation effect resulting from reactivating the spatial location representation is also not very pronounced. Advantages of looking at a critical port affect only the speed of retrieval of linguistic information and not the accessibility of linguistic information. In the general discussion, I will discuss why the effect in language production was stronger than in the experiments of the current chapter.

To conclude, the looking at nothing effect reported earlier (Richardson & Spivey, 2000) appears to be less prominent than assumed. However, the model suggested by Altmann & Kamide (2007) remains valid, and for the first time evidence for a facilitation effect predicted by Ferreira, et al. (2008) was presented in language comprehension.

Part II: Regressive eye movements in reading garden path sentences

5 Chapter 5: Introduction

The first part of this thesis investigated the role of spatial location information on language production and during retrieval of information conveyed by spoken language. Spatial location information is an integral part of the memory system, and activating it can facilitate language production and comprehension. In the second part of this thesis, I examined whether these findings also apply to reading. Do readers have a spatial representation of words within a sentence? Is spatial information utilised to support reading? Similar to the language production and comprehension experiments, I investigated whether people return to critical regions, when reprocessing information related to these regions.

A promising way to investigate the role of spatial location information in reading is to investigate regressive eye movements. As briefly introduced in Section 1.3.4, readers regress for a number of reasons such as in response to low-level visuomotor, or word identification processes or due to higher-level syntactic and semantic processes (Vitu, 2005). In this part of the thesis, I am interested in regressions initiated by higher level syntactic processes. Compared to regressions caused by low-level visuomotor processes, such regressions are usually longer and target regions beyond the word boundary (Frazier & Rayner, 1982; Meseguer, et al., 2002). If spatial location and linguistic information are tightly linked together, I predict that readers directly return to regions within a sentence, when analysing words within this region.

In the remainder of this introduction, I first summarise the literature of regressive eye movements in reading locally ambiguous sentences. For a more detailed explanation of parsing of ambiguous sentences refer to Section 1.3.2.

Furthermore, I discuss the utilisation of spatial indices in reading. I conclude this section by presenting the aims of this second part of the thesis.

5.1 Regressions during reanalysis of locally ambiguous sentences

In an early study, Frazier & Rayner (1982) suggested three possible hypotheses of how readers perform regressive eye movements in order to reanalyse a sentence; *the forward reanalysis hypothesis*, *the backwards reanalysis hypothesis*, and *the selective reanalysis hypothesis*. According to the forward reanalysis hypothesis, readers immediately return to the beginning of the sentence in order to read the whole sentence again. The backwards reanalysis hypothesis (first introduced by Kaplan, 1972) proposes that readers backtrack step by step to find potential locations of where the misanalysis happened. The selective reanalysis hypothesis suggests that readers directly return to the ambiguous region which caused the misanalysis. Frazier & Rayner (1982) pointed out that the selective reanalysis hypothesis involves no automatic processes. Rather, for each sentence, the parser has to determine the position of the ambiguous region. Frazier & Rayner (1982) found evidence for the selective reanalysis hypothesis. Regressions in garden path sentences were analysed. In the majority of trials, readers initiated a regression from the disambiguating region and regressed directly back to the ambiguous region of the sentences. The authors interpreted this as the ability of the reader to utilise linguistic information in order to directly target the ambiguous region. In the remainder of this thesis, the disambiguating region will be also called the *breakdown region* and the ambiguous region, the *target region* of a sentence.

Meseguer, et al. (2002) found similar results using Spanish sentences. The authors compared locally ambiguous sentences in which an adverbial phrase had two possible attachment sites. The adverbial phrase could either be attached to the main verb (high attachment) or to the verb of a sentential complement (low attachment). The sentences were disambiguated at the verb of the adverbial phrase. Meseguer, et al. (2002) tested whether readers had a preference to regress directly back to the main verb or to the verb of the sentential complement depending on the disambiguation. The authors found that readers tend to regress to the high attachment sites more frequently when the sentence had to be disambiguated towards a high attachment. Similar to Frazier & Rayner (1982), they concluded that an intelligent selective reanalysis process takes place.

Mitchell, Shen, Green, & Hodgson (2008) proposed a different approach: *the time-out hypothesis*. According to this hypothesis, readers do not regress in order to reread previously seen material. Rather, during reanalysis of a sentence, the forward movement is inhibited so that no new information is processed. The parser takes a time-out in order to use all available resources for the reanalysis process. Thus, the time-out hypothesis would not predicts that regressions land on the ambiguous regions of sentences.

To investigate the time-out hypothesis, Mitchell, et al. (2008) tested whether the physical layout of the linguistic material influences regressions. According to the selective reanalysis hypothesis, regressions should only depend on locations of the ambiguous region in the sentence, independent of their relative location to the rest of the sentence. In two experiments, the location of the ambiguous region and of the disambiguating word was manipulated. However, the authors found that the first regression from the disambiguating word was not affected by this manipulation. Rather, most of the first regressive eye movements targeted the word immediately to the left of the breakdown region, independent of the distance to the ambiguous region. The ambiguous region was reached in multiple successive saccades but very often was not reached at all. Thus, they concluded that regressions are not driven only by linguistic control. Since both linguistic and spatial properties influenced regressions, they suggested a hybrid model in which the regression sequences are under a rather loosely coupled control.

Mitchell, et al. (2008) point out that the finding, that most of the first regressions land on the word to the left of the disambiguating word, is a serious problem for the experiments ran by Frazier & Rayner (1982). In this study, the ambiguous region was immediately to the left of the breakdown region. Since the first regression seems to always target the region immediately to the left, independent of the distance of the intended target, it is impossible to argue that the ambiguous region was intentionally targeted. Furthermore, Mitchell, et al. (2008) discussed the data found by Meseguer, et al. (2002). They argued that the differences in material and analyses might have been responsible for the differences in the effects. Furthermore, Malsburg & Vasishth (2009) reanalysed scanpaths of the sentences from Meseguer, et al. and did not find evidence for an immediate return to the ambiguous regions of the sentences.

5.2 Spatial coding in reading

A theoretical prerequisite for the selective reanalysis hypothesis is that the readers have to have a representation of the location of the words within the sentences. Without any internal representation of word location it would be impossible to directly regress to specific words outside the perceptual span, which normally extends 12-15 characters to the right and 3-4 characters to the left of a current fixation (McConkie & Rayner, 1976; Rayner & Pollatsek, 1989; Underwood & McConkie, 1985).

Kennedy and colleagues (Kennedy, Brooks, Flynn, & Prophet, 2003; Kennedy & Murray, 1987) suggested that readers apply a spatial code to the words in a sentence and that this code can be used to directly regress to specific words.

Inhoff & Weger (2005) specifically tested how the spatial code is used to regress to leftward parts of a sentence. Subjects read sentences and received either a probe word or a question related to a single word in the previously read sentence. The authors found that regressions were generally directed towards the target but that the first regression very often undershot it. The target was only reached in further, smaller regressions. The authors argued that the first regression is guided by the spatial memory whereas in the smaller later regressions linguistic knowledge plays a larger role (Weger & Inhoff, 2007). These studies established that a spatial representation of the previously read material exists. However, regressions were exogenously generated by cueing specific words in comparison to being initiated endogenously by an internal reanalysis process which also involves a syntactic reanalysis process. The question therefore remains whether the spatial memory is used in order to syntactically reanalyse sentences or whether regressions rather follow an automatic process.

5.3 Aims of the current experiments

To summarise, the results from former studies investigating regressive eye movements, and the use of spatial coding are not very conclusive. Earlier studies showed that spatial indices for words within sentences exist and that these indices are used to reanalyse sentences and guide eye movements to critical words within sentences (Frazier & Rayner, 1982; Meseguer, et al., 2002). However, more recent studies demonstrated that although a spatial index and linguistic control exist, this information does not lead to accurately targeting regressions (Inhoff & Weger, 2005; Mitchell, et al., 2008). Rather, the regressions very often undershoot the target and only reach it in multiple successive steps.

In the current study, regressive eye movements are tested in complex garden path sentences. However, I additionally investigated whether attention moves to the left before a regression is launched. As mentioned in Section 1.3.4, in order to initiate a forward saccade in reading, attention first moves towards the target and is then followed by a saccade. This leads to an asymmetric perceptual reading span, and the eyes almost always stay within this reading span. However, the question remains whether a similar strategy is applied to regressions. Does attention move towards the left of a current fixation before a regressive eye movement is initiated? Such a strategy would lead to an extension of the perceptual reading span to the left. The reading span would reach further than four characters, as defined by former studies investigating forward reading (McConkie & Rayner, 1976; Rayner & Pollatsek, 1989; Underwood & McConkie, 1985). Evidence that the perceptual span can change its direction exists. Section 1.3.4 introduced some studies showing that the perceptual reading span is relatively flexible. Reading text in a language in which the reading direction is from right-to-left causes the reading span to extend to the left (Pollatsek, et al., 1981). Furthermore, reading English from right-left also caused the reading span to change its directional bias (Inhoff, et al., 1989).

I thus propose that attention moves to the left and that the perceptual span extends to the left before a regression is launched. At least a subpart of regressions remain within this extended perceptual reading span. Such a strategy would explain the short regressions found by Mitchell, et al. (2008) that mostly targeted the word immediately to the left of the disambiguating word.

To summarise, this second part of the thesis focused on two main questions. First, I investigated whether regressions are accompanied by a shift of attention to the left of the current fixation. I propose that information to the left supports readers to initiate a regression. In the second part, I investigated how readers regress while recovering from a misparsed sentence. In contrast to previous studies, not only the first regressive eye movements after landing on the breakdown regions were investigated but also later regressions launched after the disambiguating words were revisited.

5.4 Experiment 6

Experiment 6 investigated regressive eye movements during reading of garden path sentences. The experiment was divided into two parts. The first part examined the dimension of the perceptual reading span during regressions. The second part investigated the regression path during reanalysis of the ambiguous regions of the garden path sentences. Of most interest was whether readers are able to directly return to the ambiguous region of a sentence in order to resolve the local ambiguity of the garden path sentences.

In order to investigate the perceptual reading span, words to the left of a given fixation were masked (i.e. words were changed to non-words). Masking occurred outside of the normal perceptual reading span that usually extends to only four characters to the left of a given fixation. If the perceptual reading span extends to more than four characters when a regression is launched, reading behaviour should be affected by the masking manipulation. Progressions served as the control condition.

In the second part of this experiment, the regression path was examined. In comparison to earlier studies in which only the first regressions were investigated, the current experiment also analysed further regressions that occurred after the reader returned to the disambiguating region of the garden path sentences.

5.4.1 Method

5.4.1.1 Participants

Thirty-six participants took part in Experiment 6. The average age of the participants was 21.9 (range 18-42). All participants were native speakers of English and had normal or corrected to normal vision. Each participant was paid £6.

5.4.1.2 Material

Experimental items

Forty garden path sentences were used. These sentences had a high record of inducing regressive eye movements. They were taken from two studies (Frazier & Rayner, 1982; Pickering & Traxler, 1998) that reported a high number of regressions launched from the disambiguating region. Examples (5.1) - (5.4) illustrate sentences from these studies.

(5.1) *As the woman edited the magazine about fishing* ***amused*** *all the reporters.*

(5.2) *The criminal confessed his sins which upset kids* ***harmed*** *too many people.*

(5.3) *Since Jay always* jogs *a mile and a half* ***really*** *seems like a very short distance*

to him.

(5.4) *Yesterday, Sally <u>found out</u> <u>the answer</u> to the difficult physics problem **was** in the book.*

The ambiguous or target region of the sentences always consisted of a verb and the following noun phrase. In sentences (5.1) - (5.4), the target region is marked by underlined words. The disambiguating or breakdown region is marked by a bold font.

Sentences (5.1) and (5.2) are examples from the sentences taken from the study by Pickering & Traxler (1998). Sentences of Type(5.1) include a subordinate clause ambiguity. The noun phrase *the magazine* is initially interpreted as the direct object of the verb *edited*. However, when encountering the disambiguating verb *amused*, the sentence must be reanalysed such that *the magazine* is part of the subject of the main clause. Since the clause *the magazine* is a semantically plausible direct object of the verb *edited*, readers are usually garden-pathed. Pickering & Traxler (1998) reported that the disambiguating region triggered 20% of all regressions.

Sentences of Type (5.2) include complement-clause ambiguities. The noun phrase *his sins* is initially analysed as being the direct noun of the verb *confessed*. However, the verb *harmed* disambiguates the sentence such that *his sins* is part of the subject of a complement clause. The noun phrase *his sins* is a semantically plausible object for the main verb *confessed*, which raises the possibility that the sentence is initially incorrectly analysed. Pickering & Traxler (1998) reported that the disambiguating region triggered 32% of all regressions.

Sentences of Type (5.3) and (5.4) were taken from the study by Frazier & Rayner (1982) and represent closure and attachment sentences respectively. The initial clause of the closure sentence *a mile and a half* could be attached as an object to the verb *jogs*. Following the late closure principle (Frazier, 1978), this should be the initial analysis. However, when processing the disambiguating region *really* the reader has to reanalyse the sentence and reanalyse the noun phrase *a mile and a half* as being the subject of the following clause.

Sentences of Type (5.4) contain an ambiguous post verbal noun phrase (*the answer to the difficult problem*). According to the minimal attachment theory Frazier (1978), this noun phrase would be attached to the main verb (*was*). However, the correct analysis of the sentence is to attach the noun phrase as the subject of an embedded sentential complement to the verb. Thus, the sentence needs to be reanalysed.

Participants in the current study were British. Thus, some of the original sentences were changed to replace American expressions with British ones. For a complete list of all sentences see Appendix F.

Masked Sentences

For each sentence a second, masked version was created, in which each character was replaced by a different character that was visually similar to the original character. Descending letters were replaced by descending letters (e.g. g → *q, p* → *g*). Ascending letters were replaced by ascending letters (e.g. *t* → *l*, *k* → *h*) and baseline letters with baseline letters (e.g. *a* → *n, c* → *s*). Consonants as well as vowels were replaced by consonants resulting in masked words that consisted only of consonants. Thus, altered strings were always non-words. For a complete list of character pairs see Appendix G.

All sentences were printed in an 18 points Courier New font. The Courier New font is a non-proportional font. Thus, all letters had a fixed space, and replacing letters within a word did not change the position of words within the sentences.

Fillers

Forty non-garden path sentences were included as fillers.

5.4.1.3 Apparatus

An SR EyeLink1000 eye tracker with a sampling rate of 1000 Hz and a spatial resolution of less than 1/4 degree was used. The tracker was tower mounted and the eye movements of the right eye were tested. The stimuli were presented at a 19-inch CRT monitor running at 140Hz. The resolution of the screen was set to 1024 x 768 pixels. The Experiment Builder software developed by SR research was used to run the experiment. The analysis was done using DataViewer, developed by SR Research Ltd. and Matlab.

5.4.1.4 Procedure

Participants were seated in a soundproof chamber. The monitor was located at a distance of 90cm. Based on this distance and on the size of the font, approximately 3.3 characters per degree of visual angle were displayed.

After reading the instruction, participants were manually calibrated using a nine-point fixation stimulus. The EyeLink software validated the calibration. If validation was poor, it was repeated. The calibration procedure was repeated approximately three times during the experiment. Between items, participants had to fixate a point in the middle of the screen. This allowed the system to perform a drift correction if necessary. Furthermore, a small rectangle was shown at the upper left side of the monitor. The rectangle was at the position of the first word of the following sentence. In order to trigger the next sentence, participants had to fixate this rectangle. Thus, when the sentence appeared on the screen, the participants' eyes were already fixating the beginning of the sentence.

The garden path sentences always fit on a single line. Masking the sentences was done in the following way. When reading along, all the words starting at the second word to the left of the currently fixated word were changed to non-words. This procedure is illustrated in Example (5.5) (the word in bold face marks the currently fixated word):

(5.5)

`As the woman edited the magazine about fishing amused all the reporters.`

Bz the **woman** edited the magazine about fishing amused all the reporters.

Bz lkc woman **edited** the magazine about fishing amused all the reporters.

Bz lkc mvwnc edited **the** magazine about fishing amused all the reporters.

As mentioned in Section (1.3.4), the perceptual reading span in forward reading extends 3-4 characters to the left. The unchanged word to the left of a currently fixated word served as a buffer word to ensure that the manipulation did not occur within the normal reading span. However, if the perceptual reading span extends to more than four characters to the left before a regression is launched, the manipulation should result in different reading behaviour.

After the regression was launched and crossed the border between the fixated word and the previous word, the original version of the sentence was presented again. Thus, when the eye landed on a previous region, participants were exposed to only the correct characters.

After participants had read and understood a sentence, they were instructed to press a key on the keyboard which triggered a yes-no comprehension question (for the questions see Appendix F). The questions were answered by pressing the left (yes) or right (no) shift key on the keyboard. The questions were included to ensure that the participants read the sentences carefully and they were told that the questions were about very small details of the sentences. In order to answer the questions correctly, the ambiguous regions of the garden path sentences had to be properly resolved. The entire experiment lasted approximately 45 minutes.

5.4.1.5 Data Analysis

According to the two main research questions, the analysis of the current experiment consisted of two main parts. In the first part, I tested whether the perceptual reading span extends to the left before a regression is launched. If the perceptual reading span extends more than four characters to the left, the manipulation of the words to the left of the current fixation should affect reading behaviour. In order to measure alterations in reading behaviour, three dependent variables were tested: First, the length of the first saccade launched from the breakdown region (*first critical saccade*), second, the duration of the last fixation on the breakdown region before the first critical saccade was launched (*last breakdown region fixation*), and third, the duration of the last gaze on the breakdown region before the first critical saccade was launched (*last breakdown region gaze*). The last breakdown region gaze comprised all fixations on the breakdown region before the first critical saccade out of the breakdown region was launched.

The second part of the analysis tested how readers regressed after fixating the breakdown regions of the garden path sentences. As mentioned before, while reading a garden path sentence, readers often initiate a regressive eye movement after they landed on the breakdown region. In the analysis of the current experiment, I was most interested in these regressions. Only regressions are considered which crossed the word boundary of the breakdown region. Intra-word regressions were not included.

These regressions were subcategorised. Figure 5.1 illustrates the subtypes of regressions; the *first regression,* the *first clean regression sweep* and the *first jagged regression sweep*.

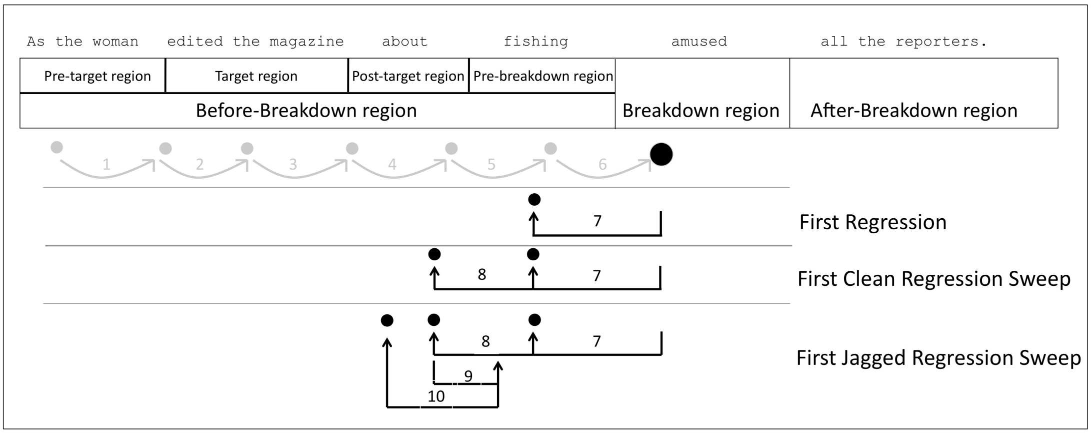

Figure 5.1: Types of regressions and regions of the garden path sentences.

The first regression denotes the first saccade launched from the breakdown region and is moving from right to left. However, in some cases, the first regression is followed by a second, or third regression. This sequence of regressions is termed a clean regression sweep. Thus, a clean regression sweep is a series of consecutive saccades moving from right-to-left (see Figure 5.1). A clean regression sweep is completed when the furthest point left of the breakdown region is reached. Another kind of regressions is the *jagged regression sweep*. A jagged regression sweep is a clean regression sweep that can be interrupted by left to right saccades within the before-breakdown region (see Figure 5.1). A current jagged regression sweep is completed when the furthest point left of the breakdown region is reached. A jagged regression sweep is divided into a first and a *second jagged regression sweep*. The first jagged regression sweep is completed when the eyes return to the breakdown or the after-breakdown region. The second jagged regression sweep starts when the before-breakdown region is visited again after the first jagged regression sweep is completed.

To analyse the landing sites of the regressions, sentences were divided into four distinct regions; the *pre-target region*, the *target region*, the *post-target region* and the *pre-breakdown region* (see Figure 5.1). The pre-target region included all words to the left of the target region, which consisted of the ambiguous verb and the following noun phrase. The region between the target region and the breakdown region was subdivided into two regions. First, the post-target region consisted of the words to the right of the ambiguous region excluding the first word left of the breakdown region. The pre-breakdown region consisted only of the first word to the left of the breakdown region. This region was treated as a separate region because Mitchell, et al. (2008) found evidence that the first regression very often targeted this region. The current experiment attempted to replicate this effect.

Most of the data was analysed by using an LME analysis (Baayen, 2008). This type of analysis was used because during the experiment there was no control over when people initiated a regression. Thus, regressions had to be selected post-hoc which resulted in an unbalanced design. LME analyses are able to deal with such unbalanced data.

5.4.2 Results

5.4.2.1 Errors

In answering the yes-no questions, participants made an error in 26.3% (378 trials) of the trials. The error rate between conditions differed only slightly; 26.6% errors were made when the sentences were not masked, and 26.0% when the sentences were masked. A t-test did not reveal significant effects ($t = .32$, $p = .75$).

5.4.2.2 Overall analysis of regressions

In relation to all saccades, 29.3% were regressions. In correctly answered trials 29.6% of all saccades were regressions and in incorrect trials 28.7% of all saccades were regressions. The breakdown region triggered 11.5% (1740) of all regressions, 10.7% (1168) in correct and 13.3% (572) in incorrect trials.

5.4.2.3 Investigation of the perceptual reading span

Two fixed-effects were compared: Masking (mask versus no-mask) and Direction (progression versus regression). The direction refers to the saccade that immediately follows the last fixation or gaze on the breakdown region. The progression & no-mask condition was used as the baseline condition (intercept). A fixed-effect of masking occurred if the progression & mask condition was significantly different from the intercept (progression & no-mask condition). A fixed-effect of direction occurred if the regression & no-mask condition showed a significant difference to the intercept. I expected no significant effect of masking in both last fixation duration, last gaze duration and saccade length, because differences in masking were analysed in the progression conditions which should remain unaffected by the masking manipulation due to the small extension of the perceptual span to the left in forward reading. I expected a small effect of direction in last fixation duration and last gaze duration due to results from former studies which showed that fixation durations preceding a regression are shorter in comparison to before a progression is launched (Altmann, Garnham, & Dennis, 1992; Rayner & Sereno, 1994). However, most crucial to the hypothesis, I expected a significant interaction between masking and direction indicating that the masking manipulation has an additional effect on the regressions but not on the progressions, as would be shown by a non-significant masking fixed-effect.

Table 5.1: The mean size of the first regression and first progression out of the breakdown region (in characters) and the mean last fixation duration and the mean gaze duration on the breakdown region just before the next saccade was launched (in ms).

Masking	Regression		Progression	
Direction	**Mask**	**No Mask**	**Mask**	**No Mask**
First critical saccade	14.9	14.8	6.4	6.2
Last breakdown region fixation	246.3	248.5	252.4	251.8
Last breakdown region gaze	291.8	297.1	302.7	300.4

In the following analysis both correct and incorrect items were included. The main interest in this part of the analysis was how the perceptual reading span changed during regressive eye movements. Thus, whether the regression led to a correct or incorrect parse is secondary. Table 5.1 shows the mean length of the first critical saccade, the mean last breakdown region fixation and the mean last breakdown region gaze.

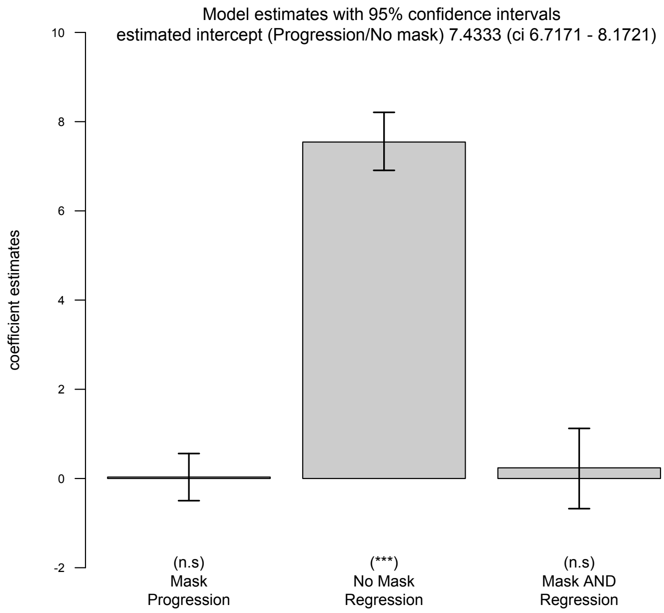

Figure 5.2: LME analysis of the length of the first saccade from the breakdown region when the region to the left of the breakdown region was masked in comparison to when it was unmasked.

Figure 5.2 illustrates the output from the LME analysis of the length of the first critical saccade. The saccade length did not differ between mask and no-mask conditions ($t = .12$, $p = .90$). A significant effect was however found in direction ($t = 23.01$, $p < .001$). Regressions were longer than progressions. No significant interaction between masking and direction was found ($t = .51$, $p = .61$). Figure 5.3 A illustrates the LME output of the analysis of the last breakdown region fixation before initiating the first critical saccade. No significant effects in Masking ($t = .27$, $p = .79$) or in Direction ($t = 1.34$, $p = .17$) were found. The Masking x Direction interaction was also not significant ($t = .02$, $p = .98$).

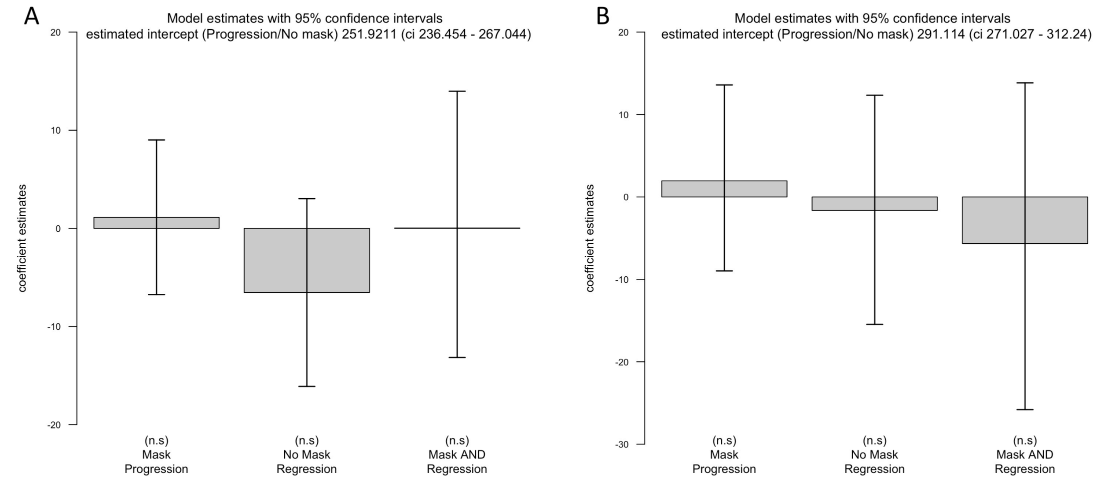

Figure 5.3 A: LME analysis of the last fixation before leaving the breakdown region when the region to the left of the breakdown region was masked in comparison to when it was unmasked.
B: LME analysis of the gaze duration before leaving the breakdown region when the region to the left of the breakdown region was masked in comparison to when it was unmasked.

Finally, Figure 5.3 B shows gaze duration on the breakdown region before initiating a saccade. Significant differences were neither found in masking ($t = .35$, $p = .72$) nor in the direction ($t = .22$, $p = .83$). The interaction between Masking and Direction was also not significant ($t = .58$, $p = .55$).

To summarise, the masking manipulation had no effect on the reading behaviour. The length of the first saccade from the breakdown region, the last fixation, and the last gaze duration on the breakdown region did not differ between masked and no masked conditions.

5.4.2.4 Analysing the regression path

Analysis of the first regression, the first clean regression sweep, the first jagged regression sweep and the second regression sweep

In this second part of the results section, I analysed how readers regressed from the breakdown region. As mentioned in Section 5.4.1.5, the first regression, the first clean regression sweep and the first and second jagged regression sweeps from the breakdown region were used as dependent variables. Only trials were included in which the breakdown region triggered a regression before the eyes moved on to further material in the post-breakdown region. By including only such trials, I could exclusively investigate regressive eye movements that were most likely triggered by the misanalysis of the garden path sentences. In total 26.2% (377 trials) of all trials showed such a pattern. In 25.4% (270 trials) of the correct trials and 28.3% (107 trials) of incorrect trials, the breakdown region triggered a regression before readers moved to a further region.

Figure 5.4 illustrates proportions of landing sites of the first regression, the first clean regression sweep, the first jagged regression sweep, and the second jagged regression sweep. Proportions of landing sites within the four regions were calculated separately for each type of regression. Thus, the sum of the proportions of one regression type in all four regions is one. A similar analysis was conducted by Mitchell, et al. (2008). Figure 5.4 A shows the regressions of the correctly answered trials and Figure 5.4 B illustrates the regressions of the incorrect trials.

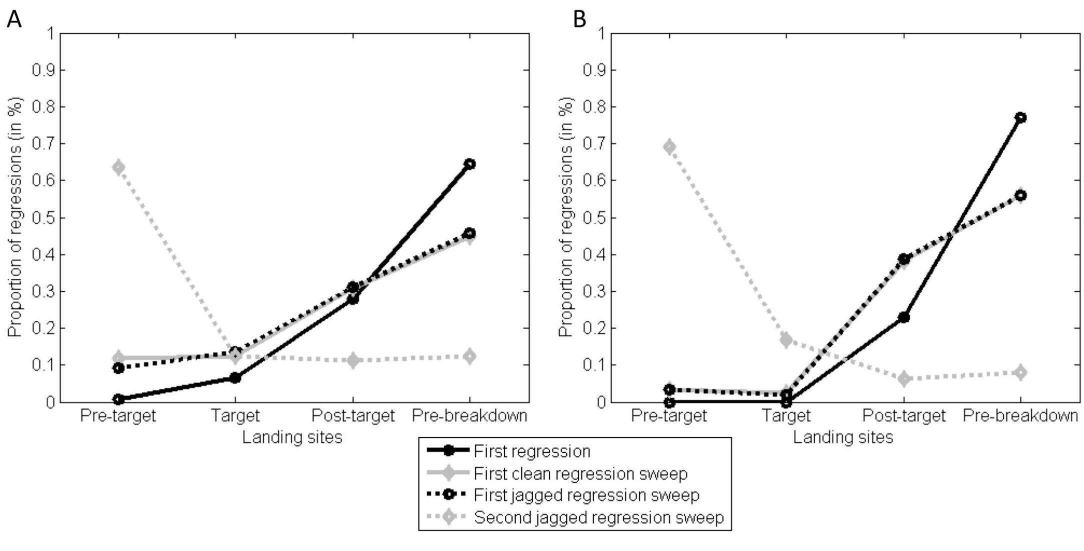

Figure 5.4 A: Proportions of landing sites of the first regression, the first clean regression sweep, the first jagged regression sweep, and the second jagged regression sweep for correct trials. B: Proportions of landing sites of the first regression, the first clean regression sweep, the first jagged regression sweep, and the second jagged regression sweep for incorrect trials.

The data was entered as categorical values into a mixed logit model (Jaeger, 2008). Mixed logit models are well suited to handle categorical outcome variables. The information whether a region was reached by a specific type of regression could be directly fed into the model. Subjects and items were entered as a random effect. Furthermore, four fixed effects were entered into the analysis: Regression Type (First regression / First clean regression sweep/ First jagged regression sweep / Second jagged regression sweep), Landing Site (pre-target region / target region/ post-target region /pre-breakdown region), Masking (no mask / mask) and Error (correct / incorrect). I was mainly interested in whether readers were able to regress directly to the target region. Thus, the target region was included in the intercept. Furthermore, since the first regression almost never reached the beginning of the sentence, it was impossible to include the first regression into the intercept. It was therefore replaced by the first clean regression sweep. Thus, the condition target region & first clean regression sweep & no mask & correct comprised the intercept

A first look at the descriptive data (Figure 5.4) reveals that it was very unlikely that regressions landed on the target region. Rather, the first regression, the first clean regression sweep, and the first jagged regression sweep mostly landed on the pre-breakdown region. Thus, the first set of regressions mostly landed on the word immediately to the left of the breakdown region. The second jagged regression sweep mostly targeted the pre-target region. Thus, after not reaching the target region in the first set of regressions, readers returned to a region at the beginning of the sentences.

Table 5.2: Fixed effects of the mixed logit model. The condition target region & first clean regression sweep & no mask & correct comprised the intercept.

Predictor	Coefficient	SE	Wald Z	p
Intercept	-2.05	0.27	7.46	<.01
Regression type: First regression	-0.45	0.43	1.045	n.s
Regression type: First jagged regression	0.00	0.39	0.00	n.s
Regression type: Second jagged regression	0.39	0.36	1.07	n.s
Landing sites: Pre-target	0.14	0.38	0.38	n.s
Landing sites: post-target	1.30	0.33	3.90	<.01
Landing sites: pre-breakdown	1.79	0.33	5.47	<.01
Masking: Mask	-1.89	1.05	1.80	n.s
Error: incorrect	0.37	0.36	1.04	n.s
Second jagged regression & pre-target	2.03	0.48	4.21	<.01
Second jagged regression & post-target	-1.93	0.51	3.80	<.01
Second jagged regression & pre-breakdown	-2.10	0.48	4.34	<.01
First Regression & pre-breakdown	1.13	0.50	2.27	<.05
Pre-breakdown & incorrect	2.46	1.10	2.24	<.05

Table 5.2 shows the output of the mixed logit model analysis. Only significant interactions were included. Since the intercept comprised the condition target region & first clean regression sweep & no mask & correct, different regression types are first compared at the target region in no masked and correct trials. The first four lines in Table 5.2 revealed no significant effects of regression types indicating that the target region was not targeted differently by the four types of regressions. The significant effect of landing sites: post-target and landing site: pre-breakdown showed that the post-target and the pre-breakdown regions were significantly more often targeted by the first clean regression sweep than the target region (target region: 13.6%, post-target region: 31.2%, pre-breakdown region: 45.9%). The effects of masking and error were not significant indicating that the first clean regression sweeps did not differ for incorrect and masked items.

Significant Regression Type × Landing Site interactions including the regression type: second jagged regression indicated that the landing sites of the second jagged regression differed from the intercept. The second jagged regression & pre target region, the second jagged regression & post-target region, and the second jagged regression & pre-breakdown region conditions different significantly from the intercept. Thus, the pre-target region was reached more often by the second jagged regression sweep (63.7%) than by the first clean regression sweep (9.3%). Furthermore, the post-target region and the pre-breakdown regions were reached less often by the second jagged regression sweep than by the first clean regression sweep (11.4%, 12.5% versus 31.2%, 45.9%). The significant first regression & pre-breakdown region effect shows that the region to the left of the breakdown region was reached more often in the first regression (64.5%) than in a first clean regression sweep (45.9%).

Except for a significant pre-breakdown region & incorrect effect, no significant effects involving the error or masking effects were found. Thus, except the pre-breakdown region, which was reached less often by the first clean regression sweep in the correct trials (45.9%) than in the incorrect trials (56.1%), no differences between items that differed in masking or error effects were found. The mask manipulation therefore, did not affect the regression behaviour nor were regressions executed differently in the correct or incorrect trials. None of the other comparisons were significant.

To summarise, the first sets of regressions mostly targeted the region adjacent to the breakdown region. The majority of the first regression, the first clean regression sweep and the first jagged regression undershot the target region and landed on a region between the target and the breakdown regions. The second jagged regression sweep mostly overshot the target region and landed on a region close to the beginning of the sentence. However, from the previous analyses it is not clear how the beginning of the sentence was reached. In the next section, I thus further investigated the size of single regressions within the second jagged regression sweep.

Size of single regressions within the second jagged regression sweep

Figure 5.5 shows the size of single regressive eye movements within the second jagged regression sweep measured in number of words (Figure 5.5 A) and number of characters (Figure 5.5 B).

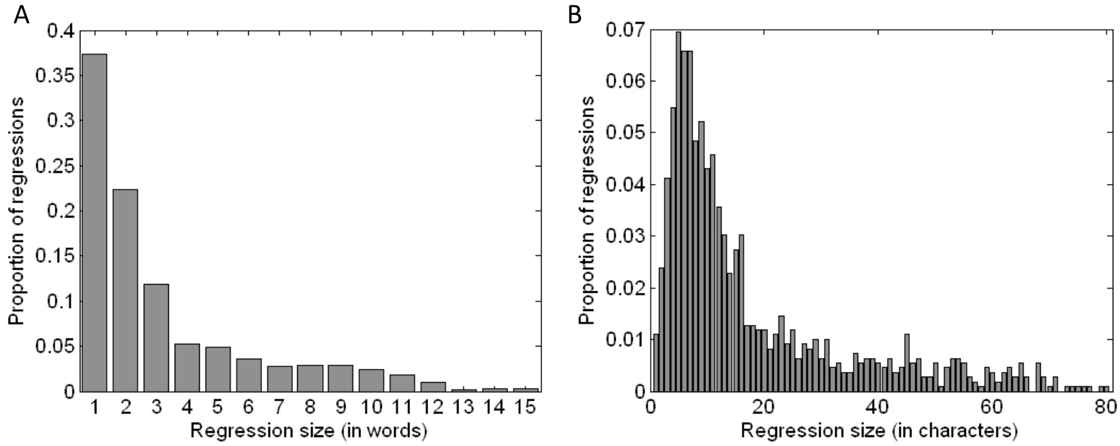

Figure 5.5 A: The size of single regressive eye movements within the second jagged regression sweep in words; B: The size of single regressive eye movements within the second jagged regression sweep in characters.

Figure 5.5 A shows that the size of the majority of regressions was only one or two words long. The analysis of the size of regressions in characters (Figure 5.5 B) supports this result. 66.8% of all regressions were shorter than 16 characters. The mean number of regressions within a second regression sweep was 2.9 which indicates that the readers mostly moved in small saccades back to the pre-target region and did not reach it by a single very long regression.

5.4.3 Discussion

The analysis of Experiment 6 was divided into two main parts. The first part tested whether readers' attention shifted to the left before a regression was launched. The second part investigated how readers initiated regressive eye movements in order to recover from a garden path sentence.

To investigate the first part, participants were presented with garden path sentences. While reading, words to the left of a given fixation were masked such that they became non-words. A buffer word was included between the current fixation and the masking manipulation. This ensured that the masked words always appeared outside the perceptual reading span which extends to 3-4 characters to the left and 12-15 to the right of a current fixation. However, if attention shifts to the left before readers initiate a regression, the perceptual reading span should be extended to more than four characters to the left and the masking manipulation becomes visible to the reader. If readers notice the manipulation, masking should alter their reading behaviour. To investigate reading behaviour, the length of the first saccade launched from the breakdown region, the duration of the last fixation, and the duration of the last gaze on the breakdown region before the first saccade was launched were measured.

No differences in reading behaviour were found when words to the left of a current fixation were masked in comparison to when they were not masked. Thus, no evidence was found that the perceptual reading span extends to the left before a regression is launched. However, an alternative interpretation of the results exists. The perceptual reading span might extend to the left before a regression is launched but reading behaviour was not changed by the masked words. Before attempting to rule out this possibility, a second experiment was conducted in order to change another possible reason for the null effects. As mention before, a buffer, consisting of one word, was included in order to prevent the masking manipulation to occur within four characters to the left. However, the size of this buffer word differed. If the word was too long, more than four characters to the left remained unchanged. Thus, the buffer might have been too long and even if the perceptual reading span expands to the left, it might not have been large enough in order to notice the masking manipulation. This issue was addressed in Experiment 7.

The second part of this experiment investigated how readers target material to the left of the breakdown region when reanalysing garden path sentences. Four different regression types were tested; the first regression from the breakdown region, the first clean regression sweep, the first jagged regression sweep, and the second jagged regression sweep. Furthermore, four different landing sites were analysed, the pre-target region, the target region, the post-target region and the pre-breakdown region. The pre-breakdown region comprised the first word to the left of the breakdown region. Of main interest was whether readers were able to land directly on the target region, which comprised the ambiguous regions of the garden path sentences.

Neither the first regression, the first clean regression sweep, nor the first jagged regression sweep often landed on the target region. Rather, participants mostly targeted the pre-breakdown region. This result replicated previous findings (Mitchell, et al., 2008) and is direct evidence against the selective reanalysis hypothesis (Frazier & Rayner, 1982) which proposes that readers are able to directly target the disambiguating region of a garden path sentence. Analyses of later occurring regressions showed that after revisiting the breakdown region, regressions targeted the beginning of the sentence with a second jagged regression sweep. This second sweep consisted of multiple smaller regressions that are mostly one or two words long.

To summarise, the results from Experiment 6 revealed no evidence that critical regions within a garden path sentence are directly targeted when such regions are reanalysed. Furthermore, no evidence for a flexible perceptual reading span was found. Thus, in reading, during reprocessing of specific information, the eyes are not immediately directed to the locations of this information. I further discuss this issue in the discussion of Experiments 6 and 7 (Section 5.6) and in the general discussion.

5.5 Experiment 7

As Experiment 6, Experiment 7 was designed to investigate the dimension of the perceptual reading span during regressions and the regression path during reanalysis of garden path sentences. Experiment 7 tested whether the perceptual reading span extends to the left before a regression is launched and how people regress after fixating the disambiguating region of a garden path sentence. However, in contrast to Experiment 6, in Experiment 7 the design was changed to address some possible problems of the experimental design.

First, in order to more closely control the amount of unmasked letters available to the left of a current fixation, the masking procedure was changed. Instead of masking all words to the left of any given fixated word, only the first and second words to the left of the breakdown region were changed when readers fixated the breakdown region. Furthermore, the buffer word was omitted. However, in order to ensure that at least four characters to the left of a given fixation remained unchanged, masking only occurred when the readers fixated the word in the breakdown region on at least the fourth character. Thus, if the perceptual reading span does not extend more than four characters to the left before a regression is launched, readers should not notice the masking manipulation.

As a second alteration in comparison to Experiment 6, the positions of the ambiguous region and the disambiguating word in the garden path sentences were kept constant. Furthermore, the length of the sentences was kept the same.

5.5.1 Method

5.5.1.1 Participants

Thirty-two participants took part in Experiment 7. The average age of the participants was 19.9 (range 18-26). All participants were native speakers of English. They had normal or corrected to normal vision. Each participant was paid £6.

5.5.1.2 Material

Experimental items

Forty garden path sentences were used in this experiment. The sentences from Experiment 6 were changed such that the ambiguous verb was always the fourth word in the sentence. The ambiguous verb was always followed by a two-word noun phrase. Like in Experiment 6, the target region always consisted of the ambiguous verb and the following noun phrase. However, the target region extended always from the fourth to the sixth word in the garden path sentences. The breakdown region was also kept constant at the eleventh word (for an example see Sentence (5.6) and the Appendix H).

(5.6) *While Mary was mending the clock in the dark hall frightened her with its chime.*

Thus, there were always four words between the last word of the target region and the breakdown region. Furthermore, the disambiguating word at the breakdown was at least nine characters long. The reason for this alteration was that it raised the probability that readers fixate the disambiguating word on at least the fourth character. The optimal viewing position on a word is just to the left of the word's centre (O'Regan, 1992) making the fourth or fifth character of a nine characters long word a likely fixation location. This was necessary because no buffer word between the breakdown region and the masked words was included in Experiment 7. Rather, the word immediately to the left of the breakdown region was masked. However, if readers fixated the word at the breakdown region on at least the fourth character, the masking manipulation should not be noticed assuming a perceptual reading span of four characters to the left of a given fixation.

The sentences were presented in 14 points Courier New font.

Masked Sentences

In order to mask the words, the characters of the words were changed in the same way as in Experiment 6.

Fillers

Forty non-garden path sentences were included as fillers.

5.5.1.3 Apparatus

The same apparatus as in Experiment 6 was used.

5.5.1.4 Procedure

The procedure of Experiment 7 was identical to Experiment 6 with the exception of how the masking manipulation was executed. Instead of masking all words to the left of a given fixation location, words were only masked when participants fixated the breakdown region. Furthermore, only the first (BR-1) and second words (BR-2) to the left of the breakdown region were masked. Example (5.7) illustrates the sentences in the four conditions. Either no word to the left of the breakdown region was changed (BR-1(no mask) & BR-2(no mask)), the first word to the left was changed (BR-1(mask)), the second word to the left was changed (BR-2(mask)), or both the first and second word to the left were masked (BR-1(mask) & BR-2(mask)).

(5.7)

`While Mary was mending the clock in the dark hall fri[`**`ghtened`**`] her with its chime.` (BR-1(no mask) & BR-2 (no mask))

`While Mary was mending the clock in the dark` *`kntt`* `fri[`**`ghtened`**`] her with its chime.` (BR-1 (mask))

`While Mary was mending the clock in the` *`bnch`* `hall fri[`**`ghtened`**`] her with its chime.` (BR-2 (mask))

`While Mary was mending the clock in the` *`bnch kntt`* `fri[`**`ghtened`**`] her with its chime.` (BR-1(mask) & BR-2 (mask))

No buffer word was included. To ensure that the masking always occurred outside the normal perceptual reading span, the words to the left were only changed when the word at the breakdown region was fixated on at least the fourth character (the bold region in Example (5.7)). Thus, since the reading span extends only four characters to the left, the masking manipulation should have no effect on forward reading. In case the reading span extends further than four characters to the left before a regression is launched, differences in reading behaviour should be observed.

After participants read the sentences, yes-no comprehension question were asked. For a list of the questions see Appendix H.

5.5.1.5 Data Analysis

The data analysis was done in the same way as in Experiment 6.

5.5.2 Results

5.5.2.1 Errors

Participants made an error in 33.9% (434 trials) of all trials. The error rate differed only slight between conditions. In condition BR-1(mask) & BR-2(no mask) 35.9% (115 trial) were answered incorrectly. In condition BR-1(no mask) & BR-2(mask) 31.1% (101 trials), in condition BR-1(mask) & BR-2(mask) 33.1% (106 trials) and in condition BR-1(no mask) & BR-2(no mask) 32.2% (103 trial) were answered incorrectly. An ANOVA analysis did not reveal any significant differences (all Fs < 2.8).

5.5.2.2 Overall analysis of regressions

With respect to all saccades, participants initiated a regression in 27.1%. In correctly answered trials 27.4% of all saccades were regressions and in incorrect trials 26.4% of all saccades were regressions. The breakdown region triggered 11.4% (1244 regressions) of all regressions, 10.6% (779 regressions) in correct and 12.8% (465 regressions) in incorrect trials.

5.5.2.3 Investigation of the perceptual reading span

In order to test the dimensions of the perceptual span, regressions and progressions that were initiated from the breakdown region were tested. To investigate the general behaviour of regressive eye movements, no distinction between correct or incorrect trials were made. However, only saccades initiated from at least the fourth character of the breakdown region were included. Thus, 6.2 % (682 regressions) of all regressions and 9.2 % (2707 progressions) of all progressions were included in the current analysis. As in Experiment 6, the first critical saccade, the last breakdown region fixation and the last breakdown region gaze were tested

Three fixed effects were included: Masking of BR-1 (BR-1(mask) versus BR-1(no mask)), Masking of BR-2 (BR-2(mask) versus BR-2(no mask)) and Direction (regression versus progression).

Table 5.3: The mean size of the first regression and first progression out of the breakdown region (in characters), the mean last fixation duration, and the mean gaze duration on the breakdown region just before the next saccade was launched (in ms).

Direction		**Regression**		
Masking	**BR-1(no mask) BR-2(no mask)**	**BR-1(mask) BR-2(no mask)**	**BR-1(no mask) BR-2(mask)**	**BR-1(mask) BR-2(mask)**
First critical saccade	18.7	14.7	17.6	13.7
Last breakdown region fixation	232.7	254.7	230.5	249.8
Last breakdown region gaze	296.0	324.7	300.7	354.0
Direction		**Progression**		
Masking	**BR-1(no mask) BR-2(no mask)**	**BR-1(mask) BR-2(no mask)**	**BR-1(no mask) BR-2(mask)**	**BR-1(mask) BR-2(mask)**
First critical saccade	7.1	6.6	7.1	7.0
Last breakdown region fixation	234.4	233.9	236.8	246.2
Last breakdown region gaze	358.2	344.5	349.1	353.4

As in Experiment 6, an LME analysis was used to investigate the effects. The condition BR-1(no mask) & BR-2(no mask) & progression was used as the baseline condition (intercept). A fixed effect of BR-1 occurred if the condition BR-1(mask) & BR-2(no mask) & progression differed from the intercept. A fixed effect of BR-2 occurred if the condition BR-1(no mask) & BR-2(mask) & progression differed from the intercept. Finally, an effect of direction occurred if the condition BR-1(no mask) & BR-2(no mask) & regression differed from the intercept. If the perceptual reading span extends to the left before a regression is launched, effects of the masking manipulation on the reading behaviour should have been measured only when readers regressed. Since the intercept included the progressions, fixed effects of BR-1 and BR-2 were only measured in the progression conditions. Thus, no effects of BR-1 and BR-2 were expected. However, a significant interaction of Direction × BR-1 and Direction × BR-2 was expected. Such interactions would indicate that masking the BR-1 and BR-2 region affected reading behaviour before a regression was launched but not before a progression was launched.

Table 5.3 illustrates the mean length of the first critical saccade launched from the breakdown region, the mean duration of the last breakdown region fixation, and the mean duration of the last breakdown region gaze. Figure 5.6 shows the output of the LME analysis for the length of the first critical saccade launched from the breakdown region.

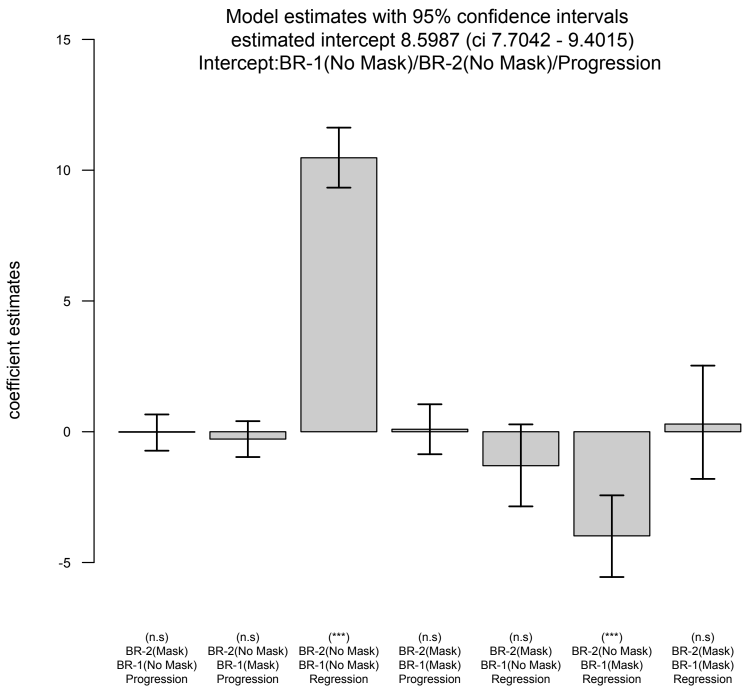

Figure 5.6: LME analysis of the length of the first saccade from the breakdown region. The saccade lengths were compared between masking of the first and second word to the left of the breakdown region and between regressions and progressions.

No fixed effects of BR-1 and BR-2 were found. Neither condition BR-1(mask) & BR-2(no mask) & progression (t = .80) nor condition BR-1(no mask) & BR-2(mask) & progression (t = .05) differed from the intercept. However, a significant effect of direction was measured. Regressions were significantly longer than progressions (t = 17.78, p < .001). Interestingly, and most crucial to our hypothesis, the condition BR-2(no mask) & BR-1(mask) & regression significantly differed from the intercept (t = 4.99, p < .001). This interaction indicates that masking the first word left of the breakdown region affected the saccade length of a regression but not of a progression. Regressions were 4.0 characters shorter when the BR-1 region was masked in comparison to when it was not masked. In contrast, the effect was only 0.5 characters when a progression was launched. No further significant effects were found.

Figure 5.7 shows the LME output of the duration of the last breakdown region fixation.

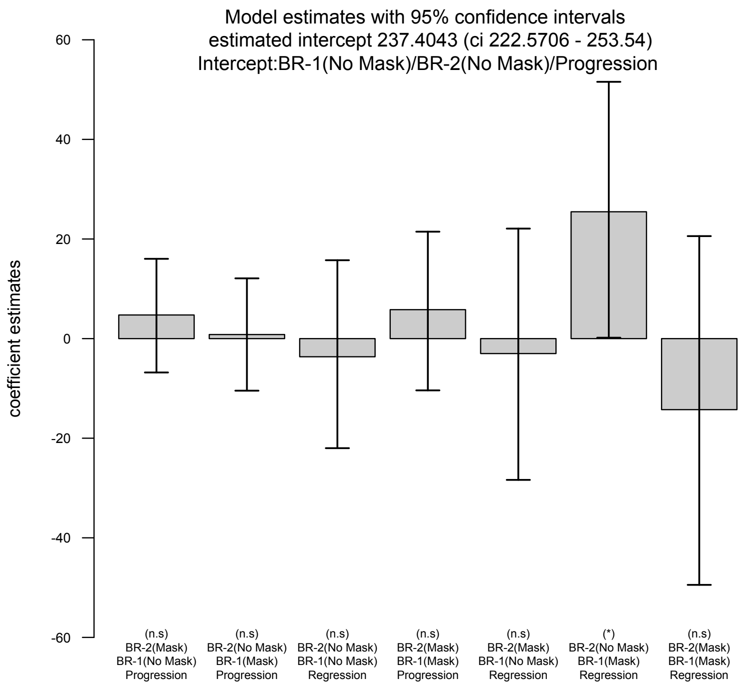

Figure 5.7: LME analysis of the last fixation before leaving the breakdown region. Fixation durations were compared between masking of the first and second words to the left of the breakdown region and between regressions and progressions.

No significant fixed effects of BR-1 ($t = .15$) or BR-2 ($t = .82$) were found. Furthermore, the effect of direction was also not significant ($t = .36$). However, similar to the analysis of the first critical saccade, condition BR-2(no mask) & BR-1(mask) & regression significantly differed from the intercept ($t = 1.94$, $p < .05$). The duration of the last fixation on the breakdown region before a regression was launched was 22ms longer when the BR-1 region was masked in comparison to when it was not masked. The difference of fixation duration between masking of region BR-1 and not masking of BR-1 was only 0.5ms when a progression was launched. All other comparisons were not significant. The output of the LME analysis of the duration of the last gaze on the breakdown region before leaving the breakdown region is depicted in Figure 5.8. No effects of BR-1 ($t = 1.08$) or BR-2 ($t = .76$) were found. Gaze durations, however, were shorter when the following saccade was a regression in comparison to a progression ($t=2.85$, $p<.01$). Furthermore, as in the analysis of last fixation the condition BR-2(no mask) & BR-1(mask) & regression significantly differed from the intercept ($t=2.15$, $p<.05$). The duration of the last gaze on the breakdown region before a regression was launched was 28.7ms longer when the BR-1 region was masked in comparison to when it was not masked. The difference in gaze duration between masking of region BR-1 and not masking of BR-1 was 13.7ms when a progression was launched. All other comparisons did not reach significance.

To summarise, masking the first word to the left of the breakdown region affected reading behaviour only when a regression was launched. In contrast to progressions, regressions were shorter when the BR-1 region was masked in comparison to when it was not masked. Similar effects were found for the duration of the last fixation and the last gaze on the breakdown region. Masking the first word to the left had only an effect when readers initiated a regression and did not affect fixation and gaze duration when a progression was launched.

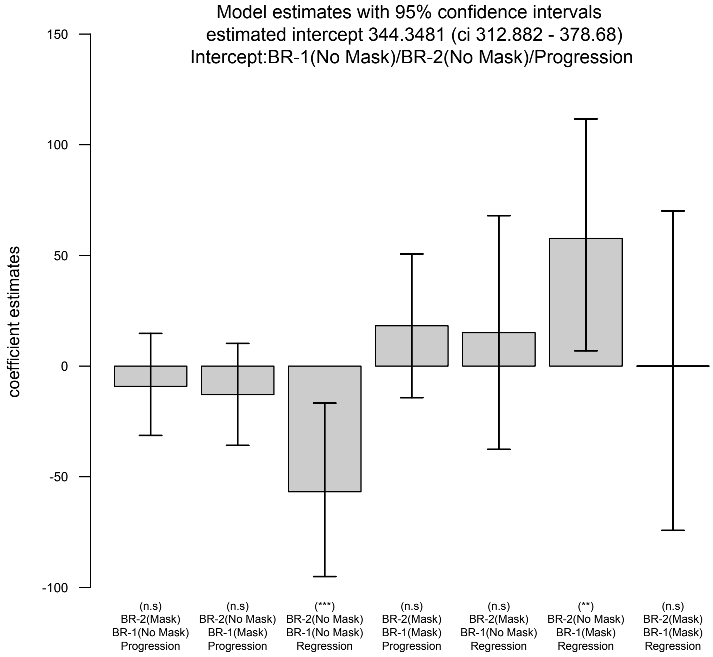

Figure 5.8: LME analysis of gaze duration before leaving the breakdown region. Gaze durations were compared between masking of the first and second words to the left of the breakdown region and between regressions and progressions.

Furthermore, effects of direction were found for the first critical saccade length and duration of the last breakdown region gaze. The length of progression was shorter in comparison to the length of a regression. Furthermore, gaze duration on the breakdown region was shorter when the following saccade was a regression in comparison to a progression.

5.5.2.4 Analysing the regression path

Analysis of the first regression, the first clean regression sweep, the first jagged regression sweep and the second regression sweep

In this second main part of the analysis, the regression path was analysed. The analysis was conducted in the same way as in Experiment 6. Only trials were included in which the breakdown region triggered a regression before the eyes moved on to further material in the post-breakdown region. In total 24.6 % (315) of all trials showed such a pattern. In 23.2 % (196) of the correct trials and 27.4 % (119) of incorrect trials, the breakdown region triggered a regression before the reader moved to further regions.

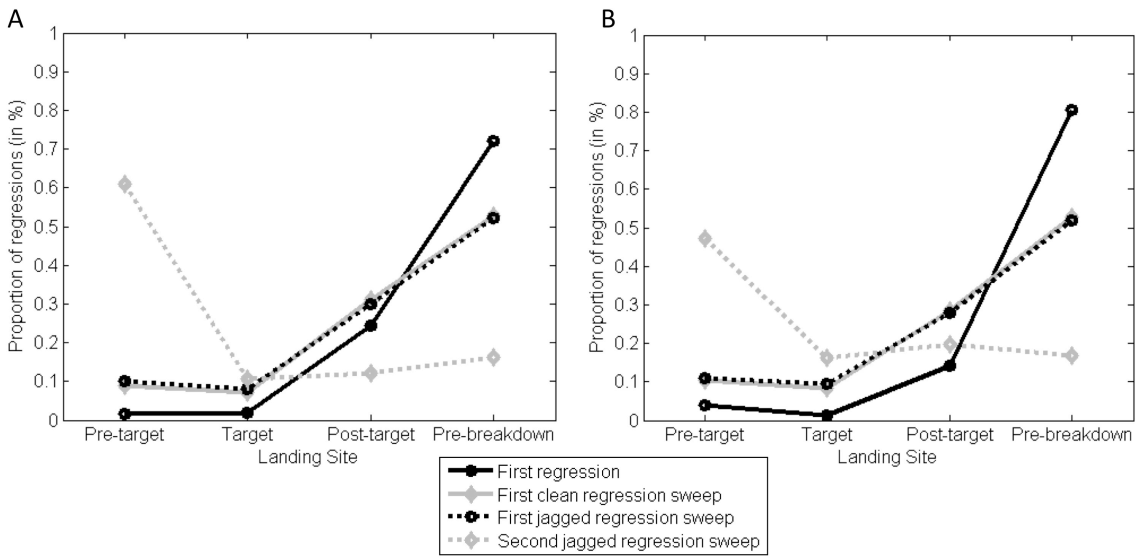

Figure 5.9 A: Proportions of landing sites of the first regression, the first clean regression sweep, the first jagged regression sweep, and the second jagged regression sweep for correct trials. B: Proportions of landing sites of the first regression, the first clean regression sweep, the first jagged regression sweep, and the second jagged regression sweep for incorrect trials.

Figure 5.9 illustrates the proportions of landing sites of the first regression, the first clean regression sweep, the first jagged regression sweep and the second jagged regression sweep. Figure 5.9 A shows the landing sites for correct trials and Figure 5.9 B shows the landing sites for incorrect trials.

The data were analysed with a mixed logit model. Three fixed effects were entered into the analysis: Regression Type (first regression / first clean regression sweep/ first jagged regression sweep / second jagged regression sweep), Landing Site (pre-target region / target region/ post-target region /pre-breakdown region), and Error (correct / incorrect). The condition target region & first clean regression sweep & correct comprised the intercept.

The results were very similar to the results of Experiment 6. As illustrated in Figure 5.9 only very few regressions landed on the target region. Rather, a high proportion of first regressions, first clean regression sweeps and first jagged regression sweeps landed on the pre-breakdown region. The second jagged regression sweep mostly targeted the pre-target region.

Table 5.4: Fixed effects of the mixed logit model. The condition target region & first clean regression sweep & correct comprised the intercept.

Predictor	Coefficient	SE	Wald Z	p
Intercept	-2.12	2.31	-9.18	p<.01
Regression type: First regression	-1.18	4.49	-2.62	p<.01
Regression type: First jagged regression	1.59	3.27	0.00	n.s
Regression type: Second jagged regression	-1.71	3.38	-0.50	n.s
Landing sites: Pre-target	1.975	3.15	0.63	n.s
Landing sites: Post-target	1.13	2.81	4.00	p<.01
Landing sites: Pre-breakdown	2.10	2.72	7.73	p<.01
Error: incorrect	-6.52	4.53	-1.44	n.s
Second jagged regression & pre-target	2.32	4.26	5.45	p<.01
Second jagged regression & post-target	-1.26	4.57	-2.75	p<.01
Second jagged regression & pre-breakdown	-1.64	4.22	-3.90	p<.01
First regression & pre-breakdown	2.11	4.97	4.25	p<.01
Second jagged regression & incorrect	1.40	5.69	2.47	p<.05
Second jagged regression & pre-target & incorrect	-1.45	7.24	-2.01	p<.05
Second jagged regression & pre-breakdown & incorrect	-1.47	6.96	-2.11	p<.05

Table 5.4 shows the output of the mixed logit model analysis. Only significant interactions were included. Since the intercept consisted of the condition target region & first clean regression sweep & correct, different regression types are first compared at the target region and for correct trials.

A significant effect of Regression type: first regression was found. The target region was less often reached by a first regression (1.8%) than by a first clean regression sweep (7.1%). No other significant effects of regression types were found. Similar to Experiment 6, the significant effect of landing sites: post-target and landing site: pre-breakdown showed that the post-target and the pre-breakdown regions were significantly more often targeted by the first clean regression sweep than the target region (target region: 7.1%, post-target region: 31.2%, pre-breakdown region: 52.9%). The fixed effect of error was not significant, indicating that the first clean regression sweeps did not differ for correct and incorrect items.

A number of significant Regression Type × Landing Site interactions were found. The second jagged regression & pre target region, the second jagged regression & post-target region, and the second jagged regression & pre-breakdown region conditions different significantly from the intercept. These interactions indicated that the pre-target region was reached more often by the second jagged regression sweep (61.1%) than by the first clean regression sweep (8.8%). Furthermore, the post-target region and the pre-breakdown regions were reached less often by the second jagged regression sweep than by the first clean regression sweep (12.2%, 16.1% versus 31.2%, 52.9%). The significant first regression & pre-breakdown region effect shows that the region to the left of the breakdown region was reached more often in the first regression (72.0%) than in a first regression sweep (52.9%).

The second jagged regression sweep differed between correct and incorrect trials. The second jagged regression & incorrect effect differed significantly from the intercept indicating that the target region was less often targeted by the second jagged regression in correct trials (10.6%) than in incorrect trials (16.4%). Significant effects of second jagged regression & pre-target & incorrect and second jagged regression & pre-breakdown & incorrect indicated that the pre-target region was more often targeted by the second jagged regression in correct trials (61.1%) than in incorrect trials (47.2%). Furthermore, the pre-breakdown region was less often targeted by the second jagged regression sweep in correct trials (16.1%) than in incorrect trials (16.9%). None of the other comparisons were significant.

In summary, the results of how readers regress from the breakdown region almost perfectly replicated the results from Experiment 6. Most of the first regressions, the first clean regression sweeps and the first jagged regression sweeps landed on the pre-breakdown region and were thus targeting the word directly adjacent to the breakdown region. The target region was almost never directly targeted. The second jagged regression sweep mostly landed near the beginning of the sentences.

Size of single regressions within the second jagged regression sweep

In the following analysis, single regressions within the second jagged regression sweep are analysed. Figure 5.10 illustrates the size of the single regressions in words (Figure 5.10 A) and in characters (Figure 5.10 B).

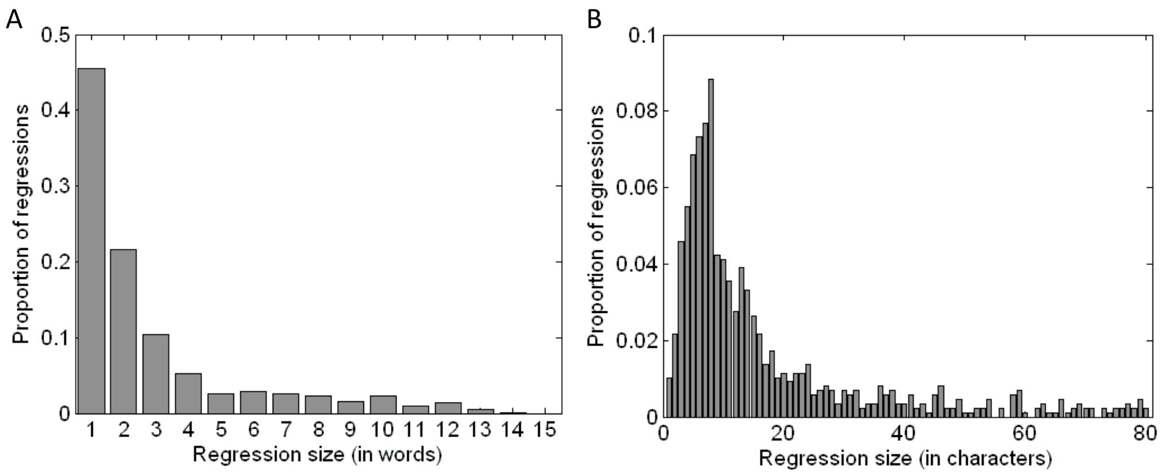

Figure 5.10 A: The size of single regressive eye movements within the second jagged regression sweep in words; B: The size of single regressive eye movements within the second jagged regression sweep in characters.

As in Experiment 6, most of the single regressions within the second jagged regression sweep were one (45.5%) or two words (21.7%) long. Figure 5.10 B shows that 68.5% of all regressions were shorter than 16 characters. Thus, in the second jagged regression sweep, readers returned to the beginning of the sentences by executing multiple smaller regressions. The mean number of regression within the second jagged regression sweep was 2.8.

A global view on the regression data

The conformity of the positions of the target and breakdown regions in the sentences of the current experiment made it possible to create a more global view on the eye movements within garden path sentences. Figure 5.11 presents the course of fixations within the garden path sentences over time. The x-axis represents the position of the words within the garden path sentences. The example sentence *While Mary was mending the clock in the dark hall frightened her with its chime* was inserted to mark target and breakdown regions. The target region includes the ambiguous verb *mending* and the following noun phrase *the clock*. The breakdown region is represented by the word *frightened*. The y-axis denotes the time course. However, instead of plotting the time in milliseconds, it was plotted in variable time slots. The reason for this procedure was that each participant read the sentences with different speed and different number of fixations. Thus, a method of normalising the reading data had to be implemented.

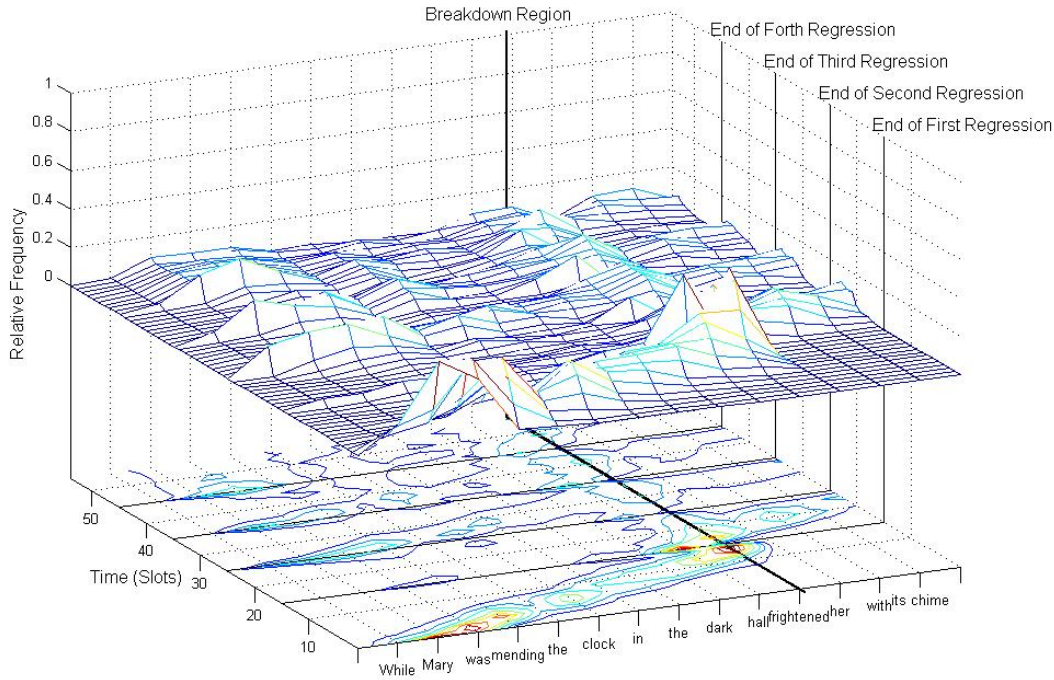

Figure 5.11: A global view on eye movements during reading garden path sentences.

At first, 15 time slots were assigned to the time it took readers to initiate the first jagged regression sweep. Thus, the time it took a participant to read a sentence to the breakdown region was divided into 15 equal time slots. The distribution of fixation locations was plotted in each of these time slots. Further

10 slots were assigned to each following jagged regression sweep. Thus, slots 15, 25, 35, and 45 represent the times, when readers regressed furthest to the left of the breakdown region. The z-axis illustrates the relative frequencies of the fixated locations.

Only items were illustrated in which readers started a regression from the breakdown region before fixating the post-breakdown region. The graph in Figure 5.11 shows how participants started reading the sentence from left to right until fixating the breakdown region. The first jagged regression sweep mostly targeted the word immediately to the left of the breakdown region. The following jagged regression sweeps returned to a region close to the beginning of the sentence. Following such a long jagged regression sweep, readers read the sentence again until they refixated the breakdown region. Presumably, if the ambiguity was not resolved, readers started another jagged regression sweep and regressed back to the beginning of the sentences. Out of 315 trials in which participants regressed from the disambiguating region without previously fixating the post-breakdown region, in 284 trials a second jagged regression sweep was initiated. A third jagged regression sweep was initiated in 195 trials. More than three jagged regression sweeps were launched in 114 trials.

5.5.3 Discussion

Experiment 7 tested the same research questions as Experiment 6. First, does the perceptual reading span extend to the left before a regression is launched? Second, how do readers regress when reanalysing a locally ambiguous sentence? Are readers able to directly target the ambiguous region of a garden path sentence?

The results from Experiment 6 suggest that the perceptual reading span does not extend further than four characters to the left before a regression is launched. However, the experimental design was changed in Experiment 7 such that the amount of unmasked information to the left was more closely controlled. The length of the first saccade from the breakdown region and the durations of the last fixation and gaze on the breakdown region were tested.

Saccade length, fixation durations, and gaze durations were compared between masked and unmasked conditions. Furthermore, influences of the masking manipulation were tested when readers initiated a regression in comparison to a progression.

The length of the first regression from the breakdown region was shorter when the first word to the left of the breakdown region was masked in comparison to when it was not masked. In contrast, masking did not influence the length of the first progression from the breakdown region. Since masking always occurred further left than four characters, readers must have been able to process information outside the normal perceptual reading span which only extends to four characters to the left. Similar results were found for the duration of the last fixation and last gaze on the breakdown region. Before initiating a regression, the duration of the last fixation and the last gaze were longer when the first word to the left was masked in comparison to when it remained unmasked. Masking did not influence fixation or gaze duration when the next saccade was a progression.

The analysis of the regression path replicated the results from Experiment 6. The first regressions, the first clean regression sweeps, and the first jagged regression sweeps all targeted the region adjacent to the breakdown region. Thus, the target region was not reached with the first set of regressions. The second jagged regression sweeps mostly landed close to the beginning of the sentences. The second sweep consisted of multiple smaller regressions. Thus, readers did not return to the beginning of the sentence by one large regression, but returned in smaller steps. Most important however, the ambiguous region was rarely directly targeted. Rather, after regressing to the word to the left of the breakdown region, readers repeatedly returned to the beginning of the sentence, in order to reread the whole sentence.

The comparison between fixation and gaze durations before the launch of a regression or progression revealed one further finding. Gaze durations on the breakdown region before a regression were shorter in comparison to gaze durations on the breakdown region before a progression. Similar results were shown in previous studies (Altmann, et al., 1992; Mitchell, et al., 2008; Rayner & Sereno, 1994) and the effect seems to be relatively stable. In comparison to Mitchell, et al. (2008), the effect was not shown in the last fixation. However, the number of replications of this effect points towards a stable phenomenon. As Mitchell, et al. (2008) mentioned, the effect is unusual, considering that a regressive eye movement is connected to more work of the parser. Furthermore, the results of the current study suggest that attention moves to the left, which might even take more time. The current study does not attempt to explain this finding but it is certainly worth investigating it more closely in the future.

5.6 General discussion of Experiments 6 and 7

In this second part of the thesis, I investigated the role of spatial location information in reading. In summary, only very little evidence for an utilisation of spatial information was found. The ambiguous target regions of the garden path sentences were almost never reached by a first regression. Even later regression types landed on the target region in just about a tenth of the trials. Instead, first regressions landed mostly on the word to the left of the breakdown region. After revisiting the breakdown or post-breakdown region, a second sweep of regressions mostly landed on a region close to the beginning of a sentence. This second jagged sweep is composed of smaller regressions that are mostly one or two words long.

In the second part of this set of experiments, the dimension of the perceptual reading span during regressions was analysed. The results from Experiment 7 suggest that before a regression is launched, the perceptual reading span extends further than four characters to the left. Readers showed differences in reading behaviour when the first word to the left was changed before a regression was launched. No differences in reading behaviour were measured when a forward saccade was launched. This result indicates that the asymmetry of the perceptual span is not statically directed towards the general reading direction of a language. Rather, it depends on the intended direction of a following saccade. Thus, attention precedes the movements of the eyes in whatever direction the eyes are about to move. Similar results were shown for people looking at a scene (Gersch, Kowler, Schnitzer, & Dosher, 2008; Smith & Henderson, 2009).

Although the results of the current study suggest that attention precedes regressive eye movements, it is unclear whether the material to the left is also lexically processed before moving the eyes, similar to parafoveal-on-fovea effects found in forward reading (Kennedy & Pynte, 2005). In the current study, I did not attempt to answer this question. However, considering the fact that readers very often regress only one or two words, and that readers are able to perceive information of one or two words to the left, information to the left of the current fixation might be used in order to control regressive eye movements.

Why did readers show no evidence for any utilisation of the spatial location representation of words in the current experiments? As mentioned in the introduction, it was shown that readers create a spatial code for words within a sentence (Kennedy, et al., 2003; Kennedy & Murray, 1987). Furthermore, Frazier & Rayner (1982) and Meseguer, et al. (2002) presented evidence for a selective reanalysis process in garden path sentences. Mitchell, et al. (2008) proposed a hybrid account that included some linguistic control, and they found some regressions that targeted the ambiguous regions.

In Section 5.1, I summarised some problems with the material of the studies by Frazier & Rayner (1982) and Meseguer, et al. (2002) that might explain why readers regressed selectively to ambiguous regions within the sentences in these experiments, and why they found evidence for the selective reanalysis hypothesis.

However, an additional factor might be responsible for the variation of the results. The material of the different studies might have varied strongly in complexity. In all studies, comprehension questions were asked after participants read the ambiguous sentences. Frazier & Rayner (1982) reported an almost perfect score of 100%. Meseguer, et al. (2002) found an average of 92% of correctly answered questions. In the ambiguous sentences presented in the experiments by Mitchell, et al. (2008), participants answered comprehension question correctly in over 90% of the trials. However, accuracy of answering questions in the current study was only at 73.7% in Experiment 6 and 66.1% in Experiment 7. Differences between masked and non-masked trials were not significant. Thus, it is possible that readers apply two different strategies in reading ambiguous sentences.

Before discussing these two strategies, I shortly talk about a prerequisite of a selective reanalysis strategy. In order to directly regress to critical words, the reader needs to first syntactically reanalyse the sentence. This stands in direct contrast to the requirements of regressions in experiments investigating the spatial code (e.g. Inhoff & Weger, 2005). In the experiments of this study, regressions were directly triggered by questions regarding specific words within the sentences. Thus, no syntactic analyses had to be performed. However, in order to return to ambiguous regions of garden path sentences, readers first have to engage in an internal syntactic reanalysis. Presumably, this process is easier when the complexity of the garden path sentences is lower. Thus, in sentences with a low complexity, readers are able to directly return to the ambiguous region because activation of the spatial location of words within the ambiguous region results in eye movements towards these regions.

However, in sentences with a higher complexity, like the material of the current experiment, the spatial location of words in the ambiguous region is not activated because an internal reanalysis was much harder to perform. Readers might attempt to first regress to a region within the perceptual reading span. Thus, a first regression target words very close to the launch region. If this strategy does not resolve the ambiguity, readers return to the beginning of the sentence in order to reread it. This strategy rather reflects the forward reanalysis hypothesis.

In summary, I did not find evidence that readers return to critical regions when linguistically processing words located at these regions. This missing effect might result from a too complex sentential structure of the experimental items. However, I found that the perceptual reading span changes its direction before a regression, and I proposed that information to the left is utilised in order to launch regressions.

In the following final part of this thesis, results from the language production, comprehension, and reading experiments are summarised. I furthermore, discuss implications for the model that integrates linguistic and visual information.

Part III: General Discussion

6 Chapter 6: General Discussion of Part I and Part II

The main aim of this thesis was to investigate the interaction of visual and linguistic information during language processing. The main focus was on the role of spatial location representation in memory. I investigated how people return to relevant locations while retrieving information related to objects located at this position, or objects that were previously located at this position. In other words, I examined how people attend to a specific location in their visual environment that was fixated during encoding of information, when retrieving this information.

In the first part of this thesis, the role of spatial location was investigated during language production and during processing of information conveyed by spoken language. Both of these processes are often executed in combination with visual information. We speak about objects in our visual environment and we listen to other people talking about objects within the visual environment. Usually a representation of such objects already exists before we speak about them or listen to someone else talking about them. For example, before uttering a sentence like "*Could you please pass me the umbrella"*, we probably have seen the umbrella and know where it is located before saying the word *umbrella*. However, people return to such objects during language processing (e.g. Meyer, et al., 1998; Potter, 1975). In five experiments I examined how people return to the location of such objects and how attending to these locations influences language and information processing.

In the second part of the thesis, I investigated how readers return to relevant information while reanalysing sentences. In Western languages like English, people read from left to right. However, when reanalysing sentences, they return to previously encountered information and initiate right to left eye movements, also called regressions. In two experiments, I investigated how accurately people initiate regressions and whether attention moves to the left before such a regression is initiated.

In the remainder of the last part of the thesis, I first summarise the results from the seven experiments conducted in the thesis. Furthermore, the implications of the results for an integrated model of visual and linguistic information are discussed. I conclude this chapter by discussing the limitations of the experiments, open questions, and possible future research.

6.1 Summary of the experiments

6.1.1 Language production and comprehension

I conducted five experiments to examine the role of spatial location in language production and during retrieval of information conveyed by spoken language. In these experiments, participants were at first presented with either objects, which they needed to describe (Experiments 1-2), or with heads conveying facts, which they needed to evaluate for their correctness (Experiments 3-5). Thus, in this first stage, information about to-be produced sentences and about to-be evaluated facts was encoded. In a second stage, this information had to be retrieved. Participants either produced the sentences or evaluated the facts. In Experiments 1 and 3 eye movements were manipulated by changing the positions of either the objects in the language production task or the heads in the fact retrieval task. Assuming that participants fixated the objects or heads (Krieger, et al., 2000; Mackworth & Morandi, 1967), I was able to investigate the effects of mismatching spatial location between encoding and retrieval of information. In Experiments 2, 4, and 5, participants were presented with an empty screen during retrieval. Thus, no visual information was present. Assuming that people fixated the empty spaces in which objects or heads were presented (e.g. Johansson, et al., 2006; Richardson & Spivey, 2000), I investigated the influence of matching spatial locations in the absence of visual information.

In language production, people fixated empty regions previously occupied by objects before naming these objects. The distribution of eye movements toward the positions of to-be produced objects was very similar in comparison to eye movements towards objects that were still visible. Furthermore, I found evidence that fixating the same position during the encoding or apprehension phase as in the retrieval or formulation phase facilitated language production. Both speech onset and articulation times, which were taken as a measure for disfluency, were facilitated when speakers fixated a matching position. However, for speech onset times a stronger facilitation effect was found.

The effects in language comprehension were not as strong as in language production. During evaluation of a fact, participants fixated empty regions that were previously occupied by the head that conveyed this fact. Such a region was more often fixated than alternative regions. However, participants showed a very strong tendency to remain at the centre of the screen. Furthermore, the effects disappeared when participants were asked to engage in a different task between encoding and retrieval of the facts.

A facilitation effect was only observed in Experiment 4, in which participants fixated an empty screen during retrieval and where no intervening task was presented. When an empty region, in which the head previously announced the critical fact, was fixated, participants were faster in evaluating the facts. However, no improvement of accuracy was observed.

The results from both the language production and comprehension experiments revealed that people fixate empty regions previously occupied by an object when processing information related to this object. Facilitation effects due to fixating a matching region were observed. However, looking at nothing and facilitation effects were stronger in language production than in language comprehension.

6.1.2 Reading

In order to investigate the role of spatial location information during reading, two experiments were conducted. Participants were instructed to read garden path sentences. Words to the left of a currently fixated word were changed to non-words. In Experiment 6, all words, except a buffer word immediately adjacent to the fixated word, were changed. In Experiment 7, a maximum of two words to the left of the breakdown region of garden path sentences were changed. By using the invisible boundary paradigm, it was ensured that the changed words always appeared outside the perceptual reading span, which is thought to extend to four characters to the left in forward reading (McConkie & Rayner, 1976; Rayner & Pollatsek, 1989; Underwood & McConkie, 1985).

I investigated to what extent readers use a spatial representation of words within a sentence in order to initiate regressions towards the target region. The target region consisted of ambiguous words within the garden path sentences. Furthermore, I examined whether the perceptual reading span changed its direction before a regression was launched.

When readers initiated a regression from the breakdown region of garden path sentences, they almost never reached the target region with first regressions. The first regression, the first regression sweep, and the first jagged regression sweep mostly targeted the region to the left of the breakdown region. Only in later regression sweeps did readers regress to regions further to the left of the breakdown region. However, no evidence was found that the ambiguous region was directly targeted by regressions. Rather, further regression sweeps landed on a region close to the beginning of the sentences.

However, evidence for a flexible perceptual reading span was found. Reading behaviour was different when the first word to the left of the breakdown region was changed in comparison to when it was not changed. However, differences in reading behaviour were only observed when readers initiated a regression. The manipulation had no effect when a progression was launched. Therefore, the perceptual reading span must have extended more than four characters to the left before a regression was launched.

To summarise, no evidence for an utilisation of the spatial location representation of words within garden path sentences was found. The size of the majority of regressions was only one word. However, regressions might be aided by a flexible perceptual reading span such that attention moved to the left before a regression is launched.

In the next sections, I outline the implications of these results for an integrated model of vision and language and discuss the differences found in language production, comprehension, and reading.

6.2 The representation of spatial location in language processing

Objects in our environment continuously change their position, orientation, or visual features due to movements of the viewer, of the objects, change of lighting conditions and so forth. However, humans are perfectly able to retain a constant representation of objects. Thus, even if an object changed its visual features, it is still recognised as being the same object. How are we able to do this so effortlessly?

Kahneman, et al. (1992) proposed a model in which visual and spatial location features are integrated into a so-called object file. As summarised in Section 2.2.2.1, the spatial location feature plays a special role in retrieving or updating other features within the object file. In order to keep a constant mental representation of an object, despite its changing visual characteristics, features of an object are updated, based on its location. In other words, as long as an object does not abruptly change its position, even if it changes its visual features, we interpret it as being the same object.

Recently, this model was extended to also integrate linguistic information (Altmann & Kamide, 2007; Ferreira, et al., 2008). According to these models, spatial location, visual, and linguistic features are closely intertwined with each other. Ferreira, et al. (2008) hypothesised that such architecture might lead to a cascading activation of the features. If the linguistic feature of an object is activated, the spatial location feature is activated as well. This process leads to an eye movement towards the position of the object. Furthermore, fixating the position of the object activates the linguistic feature which leads to a facilitation of language processing.

The present study provides evidence for such a model. Activating the spatial location feature by gazing at the position of a critical object facilitates both language production and comprehension processes related to this object. Especially in language production, this effect can be observed when people fixate a position that was previously occupied by an object, or at the object itself. Thus, spatial location is an integral part of an internal, multi-featured representation of objects in memory.

In contrast to this model is the external memory account (O'Regan & Nöe, 2001). This account proposes that the outside world serves as its own memory and only very sparse internal representations are built. The information in the outside world is accessed by spatial indices that point toward information in the external environment. Such a model would not predict facilitation effects as presented in the current experiments. If no internal representation of objects exists, a repeated fixation on an empty location at which an object was located could not have boosted retrieval of information related to this object. Thus, the results of this study are evidence against such a strong external memory account.

In the next section, the differences between effects found in language production, comprehension, and reading are discussed.

6.2.1 Why do the effects differ across the three sets of experiments?

In the previous section, I concluded that visual and linguistic information are tightly integrated and that spatial location information plays an important role in retrieving linguistic information.

However, the strength of the effects varied in the language production and comprehension experiments. In the reading experiment, no evidence for an utilisation of the spatial location representation of words within sentences was found. People did not regress to critical regions during the reanalysis of garden path sentences. In the following section, I discuss what might have caused these differences.

First of all, the methodologies of the experiments were very different. In language production, participants had to name single objects only. Furthermore, the spatial relation of the objects played an important role in these experiments. In the language comprehension experiments, participants engaged in evaluating facts and in the reading experiments, people read garden path sentences. These differences in methods impede the comparability of results across experiments. Furthermore, in the reading experiments, since readers almost never directly returned to the target regions, no comparison between trials, in which people fixated critical regions and trials where they did not fixate this region, could be conducted.

Second, the level of complexity of the tasks in the three sets of experiments differed. In the language production experiments, participants only needed to remember the position of two objects. In the language comprehension experiments, people needed to remember four facts. In reading, the spatial location of more than ten words had to be memorized. Thus, it is possible and likely, that the number of objects that had to be stored impaired the ability to utilise the spatial location features. To my knowledge, there is no specific research about the number of object files that can be kept in memory and how an increasing number of object files impairs the interaction between language and vision. However, Ferreira, Engelhardt, & Jones (2009) conducted an experiment similar to previous visual world experiments (Tanenhaus, et al., 1995). In contrast to the Tanenhaus, et al. experiment, in which the number of visible objects was restricted to four, twelve objects were shown. Ferreira, et al. (2009) found no evidence that the visual context was utilised in order to disambiguate the temporal ambiguities in the presented garden path sentences. Tanenhaus, et al, (1995) however found such an effect. Ferreira, et al. (2009) concluded that a more complex visual scene diminishes the utilisation of the visual context in order to disambiguate garden path sentences.

Furthermore, Experiments 4 and 5 showed that task complexity directly influences the looking at nothing and facilitation effects. When participants engaged in an additional task between retrieval and encoding of information, no effects were found.

Further research needs to determine the exact influence of the complexity of visual scenes or the task, but the results of the experiments of the current thesis suggest that a close relationship between these factors and the looking at nothing and facilitation effects exist.

6.2.2 The use of spatial location in child development

The results of the current study suggest that the interaction with the visual environment during language production and comprehension is governed by a similar mechanism. Consequently, speakers fixate objects before talking about them and listeners fixate objects after hearing someone talk about them. When learning a language, infants seem to use a similar mechanism. It was shown that infants follow the gaze of their mother in order to link words with objects in their visual environment and that the resulting shared attention facilitates language acquisition (Baldwin, 1991, 1993; Scaife & Bruner, 1975; Tomasello, 1988, 1995). Language learning is certainly not a trivial task and infants need some linkage between words and the objects around them. Studies investigating eye movements in language production often hypothesise that name-related gazes are an important part of this link (Bock, et al., 2003; Hanna & Brennan, 2007).

The results of the current study suggest that following a speaker's eye movements towards an object even facilitates retrieval of linguistic information related to this object. Thus, in early life, infants utilise speakers' eye movements in order to facilitate word learning. Later on, these eye movements might be used in order to support the retrieval of words from memory.

However, it is unlikely that speakers consciously fixate objects longer than necessary in order to aid the comprehension process of the listener. Instead, eye movements towards objects are a consequence of the architecture of the cognitive system that supports language processing for both the speaker and the listener.

6.3 Limitations of the experiments and open questions

In the current section, I discuss some of the limitations of the experiments conducted in this study and their implications on the interpretation of the results.

6.3.1 Controlling eye movements

The language production and comprehension experiments used two different methods in order to compare eye movements that are located at position A with eye movements located at position B. The first method guided the eyes to different locations by presenting an object on an otherwise empty screen. It has been shown that people rather fixate objects on the screen instead of looking at an empty space. Thus, by changing the location of the objects, eye movements towards different locations could be compared with each other. However, by using this method, the role of the spatial location representation could not be investigated in isolation. Visual information was always present as well and might have interfered with effects resulting from the activated spatial location representation.

The second method investigated the role of spatial location by presenting empty space. The current study as well as previous studies showed that people fixate empty regions previously occupied by an object when processing information related to this object. However, in order to test whether fixating a critical empty region affects language processing, items had to be divided post hoc into items in which the critical region was fixated and items in which it was not fixated. This method does not take into account why people fixated the empty region in some trials and not in others.

Since we cannot force participants to move their eyes to certain positions, either of these methods has to be used. However, by triggering fixations by a visual stimulus, new visual information is added. By presenting no visual input, we cannot say for sure what exactly triggered the eye movement.

An alternative, less obtrusive method could be to present visual cues for a very brief time in order to trigger fixations towards a specific location. However, even in such a case, new visual information would be presented.

6.3.2 Causation of the effect

Since eye movements were not directly and unobtrusively controlled, it is difficult to determine the causation of the effect. Did people fixate empty critical regions in order to aid the retrieval of information, or was the empty region fixated because people retrieved related information. The present experiments do not allow to distinguish between these two alternatives. However, I showed that both effects are interrelated; as put forward by Ferreira, et al. (2008):

> … whether the looks are intentional or are unconsciously triggered, the conclusion is the same: looking at nothing is an entirely expected consequence of human cognitive architecture (Ferreira, et al., 2008, p. 409).

6.3.3 Facilitation or inhibition

Another open question was whether fixating a critical empty region facilitates language processing or whether fixating an alternative region inhibits language processing. For example, in the language production experiments, fixating a region in which a filler object (i.e. the typewriter) was located might have interfered with the language production process. Given that only very few participants fixated the filler objects, this hypothesis could not be tested with the present data.

6.4 Back to the pub

In the introduction I asked you to imagine a situation in which you meet an unknown person at the local pub. After being introduced to that person, you could only remember the name after re-entering the room in which that person was sitting. Some of the experiments reported here empirically tested this effect. The results suggest that the person does not need to be sitting in the usual place in order for you to conjure up his name. Just seeing the person's face might be enough. If the person is not in the pub, looking at the same place, in which you saw him before, might help to remember his name. However, the boost to the name-recall powers may be short lived and it may be a good idea to write down his name before heading to the bar.

7 References

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition, 93*(2), B79-87.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31*, 685-712.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264.

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*, 502-518.

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191-238.

Baayen, H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge University Press.

Baldwin, D. (1991). Infants'contribution to the achievement of joint reference. *Child Development, 62*, 875-890.

Baldwin, D. (1993). Early Referential Understanding: Infants' ability to recognize referantial acts for what they are. *Developmental Psychology, 29*, 832-843.

Ballard, D., Hayhoe, M., Pook, P., & Rao, R. (1997). Deictic codes for the embodiment of cognition. *Behavioral Brain Sciences, 20*, 723-767.

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and brain sciences, 22*(4), 577-609; discussion 610-560.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-362). New York: Wiley.

Biederman, I. (1972). Perceiving real world scenes. *Science, 177*, 77-80.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143-177.

Bierwisch, M., & Schreuder, R. (1992). From concepts to lexical items. *Cognition, 42*, 23-60.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language, 48*, 653-685.

Boersma, P. W., & Weenick, D. (1992-2008). *Praat: doing phonetics by computer (Version 5.0.38)*.

Boland, J. E., Tanenhaus, M. K., & Garney, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language, 29*, 413-432.

Carver, R. P. (1990). *Reading rate: A review of research and theory*. San Diego, CA, US: Academic Press.

Castelhano, M. S., & Henderson, J. M. (2005). Incidental visual memory for objects in scenes. *Visual Cognition, 12*, 1017-1040.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368-407.

Clark, V. P., Fan, S., & Hillyard, S. A. (1995). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping, 2*, 170-187.

Cooper, R. M. (1974). The Control of Eye Fixation by the Meaning of Spoken Language. *Cognitive Psychology, 6*, 84-107.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory & Language, 14*, 113-121.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164-203.

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences, 12*(11), 405-410.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic parsing. *Journal of Memory and Language, 25*, 348-368.

Ferreira, F., Engelhardt, P. E., & Jones, M. W. (2009). *Good enough language processing: A satisficing approach.* Paper presented at the conference of the Cognitive Science Society, Austin, TX.

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movement and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 555-568.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. A., Garrettt, M., Walker, E., & Parkes, C. (1980). Against definitions. *Cognition, 8*, 263-367.

Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing. *Experimental Brain Research, 142*(1), 139-150.

Frazier, L. (1978). On comprehensing sentences: Syntactic parsing strategies. Unpublished doctoral dissertation, University of Conneticut, Storrs.

Frazier, L., & Rayner, K. (1982). Making and Correcting Errors during Sentence Comprehension: Eye Movements in the Analysis of Structurally Ambiguous Sentences. *Cognitive Psychology, 14*, 178-210.

Gersch, T. M., Kowler, E., Schnitzer, B. S., & Dosher, B. A. (2008). Attention during sequences of saccades along marked and memorized paths. *Vision Research, 8*(16), 1-24.

Glenberg, A., Schroeder, J., & Robertson, D. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory and Cognition, 26*, 651-658.

Goldman, N. (1975). Conceptual generation. In R. Schrank (Ed.), *Conceptual information*. North-Holland.

Griffin, Z. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*(1), B1-14.

Griffin, Z. (2004a). The eyes are right when the mouth is wrong. *Psychological science : a journal of the American Psychological Society / APS, 15*(12), 814-821.

Griffin, Z. (2004b). Why look? Reasons for eye movements related to language production. In J. M. Henderson & F. Ferreira (Eds.), *The integration of language, vision, and action: Eye movements and the visual world.* (pp. 213-247). New York: Psychology Press.

Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological science : a journal of the American Psychological Society / APS, 11*(4), 274-279.

Griffin, Z., & Oppenheimer, D. M. (2006). Speakers gaze at objects while preparing intentionally inaccurate labels for them. *Journal of experimental psychology Learning, memory, and cognition, 32*(4), 943-948.

Hanna, J. E., & Brennan, S., E (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language, 57*, 596 - 615.

Henderson, J. M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology-General, 123*(4), 410-426.

Henderson, J. M. (2008). Eye movements and visual memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 87-121). Oxford: University Press.

Henderson, J. M., & Anes, M. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Exerimental Psychology: Human Perception and Performance, 20*, 826-839.

Henderson, J. M., & Castelhano, M. S. (2005). Eye movements and visual memory for scenes. In G. Underwood (Ed.), *Cognitive Processes in Eye Guidance* (pp. 213-235). Oxford: University Press.

Henderson, J. M., & Ferreira, F. (2004). *The interface of language, vision, and action: Eye movements in the visual world.* New York: Psychology Press.

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics, 45*(3), 196-208.

Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology-Human Perception and Performance, 29*(2), 388-402.

Hollingworth, A. (2005). Memory for object position in natural scenes. *Visual Cognition, 12*(6), 1003-1016.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of experimental psychology Learning, memory, and cognition, 32*(1), 58-69.

Hollingworth, A. (2008). Visual memory for natural scenes. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 123-162). Oxford: University Press.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Exerimental Psychology: Human Perception and Performance, 28*(1), 113-136.

Hoover, M., & Richardson, D. (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition, 108*(2), 533-542.

Huettig, F., & Altmann, G. T. M. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15*(8), 1-34.

Inhoff, A., Pollatsek, A., Posner, M., & Rayner, K. (1989). Covert attention and eye movements during reading. *The Quarterly Journal of Experimental Psychology, 41 A*(1), 63-89.

Inhoff, A., & Weger, U. W. (2005). Memory for word location during reading: eye movements to previously read words are spatially selective but not precise. *Memory & cognition, 33*(3), 447-461.

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of Language, Vision, and Action: Eye Movements and the Visual World* (pp. 105-133). New York: Psychology Press.

Jackendoff, R. (2002). *Foundation of language*. Oxford: Oxford University Press.

Jaeger, T. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and toward logit mixed models. *Journal of Exerimental Psychology, 59*, 434-446.

Jescheniak, J., & Levelt, W. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning Memory and Cognition, 20*(4), 824-843.

Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and Spoken Descriptions Elicit Similar Eye Movements During Mental Imagery, Both in Light and in Complete Darkness. *Cognitive Sc.: A Multidisciplinary J., 30*(6), 1053-1079.

Johnson, C. J. (1992). Cognitive components of naming in children: effects of referential uncertainty and stimulus realism. *Journal of Experimental Child Psychology, 53*, 24-44.

Jones, M. W., Obregón, M., Kelly, M. L., & Branigan, H. P. (2008). Elucidating the component processes involved in dyslexic and non-dyslexic reading fluency: Evidence from eye-tracking. *Cognition*, 19.

Just, M. A., & Carpenter, P. A. (1976). Eye Fixations and Cognitive Processes *Cognitive Psychology, 8*, 441-480.

Just, M. A., & Carpenter, P. A. (1978). Inference processed during reading: Reflection from eye fixations. In J. W. Sender, D. F. Fisher & R. A. Monty (Eds.), *Eye movements and the higher psychological functions*. Hillsdale, N.J.: Erlbaum.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixation to comprehension. *Psychological Review, 87*(4), 329-354.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175-219.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*, 133-156.

Kaplan, R. (1972). Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence, 3*, 77-100.

Keller, F., Gunasekharan, S., Mayo, N., & Corley, M. (2009). Timing Accuracy of Web Experiments: A Case Study Using the WebExp Software Package. *Behavior Research Methods, 41*(1), 1-12.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 14*, 185-209.

Kennedy, A. (1992). The spatial coding hypothesis. In K.Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 379-396). New York: Springer.

Kennedy, A., Brooks, R., Flynn, L., & Prophet, C. (2003). The reader's spatial code. In J. Hyöna, R. Radach & H.Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 193-212). Oxford: Elsevier.

Kennedy, A., & Murray, W. S. (1987). Spatial Coodinates and Reading: Comments on Monk (1985). *The Quarterly Journal of Experimental Psychology, 39A*, 649-656.

Kennedy, A., Murray, W. S., Jennings, F., & Reid, C. (1989). Parsing complements: Comments on the generality of the promciples of minimal attachment. *Language and Cognitive Processes, 4*, 51-76.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153-168.

Kintsch, W., & Van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological review, 85*, 363-394.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temorary syntactic closure ambuguity. *Journal of Memory & Language, 40*, 153-194.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition, 95*(1), 95-127.

Knoeferle, P., Habets, B., Crocker, M. W., & Muente, T. F. (2008). Visual scenes trigger immediate syntactic reanalysis: Evidence from ERPs during situated spoken comprehension. *Cerebral Cortex, 18*(4), 789-795.

Krieger, G., Rentschler, I., Hauske, G., Schill, K., & Zetzsche, C. (2000). Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics. *Spatial Vision, 13*(2-3), 201-214.

La Heij, W., van der Heijden, A. H. C., & Plooij, P. (2001). A paradoxical exposure-duration effect in the stroop task: Temporal segregation between stimulus attributes facilitates selection. *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 622-632.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Exerimental Psychology, 99*, 199-208.

Lachman, R., & Lachman, J. L. (1980). Picture naming: retrieval and activation of long-term memory. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg & L. W. Thompson (Eds.), *New directions in memory and aging* (pp. 313-343). Hillsdale, NJ: Erlbaum.

Laeng, B., & Teodorescu, D. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26*, 207-231.

Levelt, W. J. M. (1989). *Speaking: From Intention To Articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-75.

Mackworth, N., & Morandi, A. (1967). Gaze selects informative details within pictures. *Perception and Psychophysics, 2*(11), 547-547.

Malsburg, T., & Vasishth, S. (2009). *Individual differences in scan-paths and reanalysis strategies while reading temporarily ambiguous sentences.*

Paper presented at the European Conference for Eye Movements (ECEM), Southampton, UK.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29-63.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychol Bull, 81*(12), 899-917.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & cognition, 32*(8), 1389-1400.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Percept Psychophys, 17*(6), 578-586.

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8*(5), 365-368.

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics, 25*, 221-224.

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & cognition, 30*(4), 551-561.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition, 66*(2), B25-33.

Milne, R. (1982). Predicting garden path sentences. *Cognitive Science, 6*, 349-373.

Mitchell, D. C. (1987). Reading and syntactic analysis. In J. Beech & A. Colley (Eds.), *Cognitive approaches to reading* (pp. 87-112). Chichester, UK: John Wiley & Son Ltd.

Mitchell, D. C., Corley, M., & Garnham, A. (1992). Effects of context in human sentence parsing - Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(1), 69-88.

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the selective reanalysis hypoythesis. *Journal of Memory & Language, 59*, 266-293.

Morrison, R. E. (1984). Manipulation of Stimulus Onset Delay in Reading: Evidence for Prallel Programming of Saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 667-682.

Morton, J. (1969). Interaction of information in word recognition. *Psychological review, 76*(2), 165-178.

Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E., & Pernier, J. (2000). Neurophysiological correlates of face gender processing in humans. *European Journal of Neuroscience, 12*(1), 303-310.

Myung, J., Blumstein, S., & Sedivy, J. (2006). Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition, 98*, 223-243.

Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34*, 191-243.

O'Regan, J. (1992). Optimal viewing position in words and strategy-tactics theory of eye movements in reading. In K.Rayner (Ed.), *Eye movements and visual cognition:Scene perception and reading* (pp. 333-354). New York: Springer.

O'Regan, J., & Nöe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral Brain Sciences, 24*(5), 939-1031.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*, 273-281.

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Verbal Learning and Verbal Behaviour, 24*, 59-88.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics, 16*(2), 283–290.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and Recovery From Garden Paths: An Eye-Tracking Study. *Journal of Exerimental Psychology: Learning, Memory, and Cognition, 24*(4), 940-961.

Pollatsek, A., Bolozky, S., Well, A., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*, 174-180.

Pollatsek, A., Rayner, K., & Henderson, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 199-210.

Potter, M. C. (1975). Meaning in visual search. *Science, 187*(4180), 965-966.

Potter, M. C., Kroll, J. F., & Harris, C. (1980). Comprehension and memory in rapid, sequential reading,. In R. S. Nickerson (Ed.), *Attention and performance* (Vol. 8). Hillsdale, NJ: Erlbaum.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*, 65-81.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206*(468-469).

Rayner, K., McConkie, G. W., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 529-544.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Hove, UK: Lawrence Erlbaum.

Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: on the use of regression-contingent analyses. *Memory & Cognition, 22*(3), 281-285.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics, 9*, 912-922.

Rensink, R., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*, 368-373.

Richardson, D., Altmann, G. T. M., Spivey, M. J., & Hoover, M. (2009). Much ado about eye movements to nothing: a response to Ferreira et al.: Taking a new look at looking. *Trends in Cognitive Sciences, 13*(6), 235-236.

Richardson, D., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition, 102*(1), 129-138.

Richardson, D., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition, 76*(3), 269-295.

Scaife, M., & Bruner, J. S. (1975). The Capacity for joint visual attention in the infant. *Nature, 253*, 265-266.

Schafer, A., Carlson, K., Clifton, C., & Frazier, L. (2000). Focus and the interpretation of pitch accent: Disambiguating embedded questions. *Language and Speech, 43*(75-105).

Simons, D., & Levin, D. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review, 5*, 644-649.

Smith, T. J., & Henderson, J. M. (2009). Facilitation of return during scene viewing. *Visual Cognition, 16*(6), 1083-1108.

Snodgrass, J. G., & Vanderwart, M. (1980). A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174-215.

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behaviour Research Methods, Instruments, & Computers, 28*, 516-536.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological research, 65*(4), 235-241.

Spivey, M. J., Richardson, D., & Fitneva, S. (2004). Thinking outside the brain: Spatial indices to visual and linguistic information. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the viual world* (pp. 161-189). New York: Psychology Press.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 121*(153-156).

Tanenhaus, M. K., Carlson, G., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes, 4*, 211-234.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632-1634.

Tomasello, M. (1988). The role of joint attentional processes in early language development. *Language Sciences, 10*, 69-88.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development*: Lawrence Erlbaum Assiciates.

Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39*, 558-592.

Trueswell, J., & Tanenhaus, M. K. (1994). Toward a lexical framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155-179). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20*(2), 153-162.

Van Gompel, R. P. G., & Pickering, M. J. (2001a). Lexical guidance in sentence processing: a note on Adams, Clifton, and Mitchell (1998). *Psychonomic Bulletin and Review, 8*(4), 851-857.

Van Gompel, R. P. G., & Pickering, M. J. (2001b). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory & Language, 45*(2), 225-258.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2000). Unrestricted race: A new model of syntactic ambiguity resolution. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 621-648). Oxford: Elsevier.

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience, 13*(4), 454-461.

Vitu, F. (2005). Visual extraction processes and regressive saccades in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 1–32). Oxford, NY: Oxford University Press.

Vitu, F., & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In A.Kennedy, R.Radach, D.Heller & J.Pynte (Eds.), *Reading as a perceptual process.* (pp. 301-326). Oxford: Elsevier.

Weger, U., & Inhoff, A. (2007). Long-range regressions to previously read words are guided by spatial and verbal memory. *Memory & Cognition, 35*(6), 1293-1306.

Wheeldon, I. R., Meyer, A., & Meulen, F. (2007). Speech-to-gaze alignment in anticipation errors. In R. P. G. Van Gompel, M. H. Fischer, W. S. Murray & R. Hill, L. (Eds.), *Eye Movements: A window on Mind and Brain*. Oxford, UK: Elsevier.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehension mentally represent the shape of objects. *Psychological Science, 13*(168-171).

A. Appendix: Experiment 1 - Experimental items

Objects were selected from the Hemera® Photo-Objects collection or downloaded from Internet sources. The COBUILD spoken logarithmic frequency (CobSLog) was obtained from the CELEX lexical database.

Item	Objects name	Number of letters	Number of syllables	CobSLog frequency
1	saw	3	1	0.00
1	dice	4	1	0.00
1	harp	4	1	0.00
2	statue	6	2	0.60
2	orange	6	2	0.60
2	razor	5	2	0.70
3	banana	6	3	0.70
3	umbrella	8	3	1.04
3	typewriter	10	3	1.04
4	skirt	5	1	1.08
4	bone	4	1	1.11
4	plug	4	1	1.20
5	coat	4	1	1.20
5	pen	3	1	1.28
5	shell	5	1	1.32
6	bullet	6	2	0.60
6	compass	7	2	0.60
6	feather	7	2	0.60
7	earring	7	2	0.00
7	goggles	7	2	0.00
7	bunk bed	8	2	0.00
8	wheel chair	11	2	0.00
8	robot	5	2	0.30
8	medal	5	2	0.30
9	suitcase	8	2	0.30
9	dummy	5	2	0.30

9	kettle	6	2	0.48
10	tomato	6	3	0.48
10	pyramid	7	3	0.70
10	diamond	7	3	0.70
11	tooth	5	1	1.20
11	brick	5	1	1.28
11	ski	3	1	1.48
12	hand	4	1	2.49
12	car	3	1	2.56
12	book	4	1	2.88

B. Appendix: Experiment 2 - Experimental items

Objects were selected from the Hemera® Photo-Objects collection or downloaded from Internet sources. The COBUILD spoken logarithmic frequency (CobSLog) was obtained from the CELEX lexical database.

Item	Objects name	Number of syllables	H-Value	CobSLog frequency
1	bin	1	0.00	0.00
1	saw	1	0.00	0.00
1	rope	1	0.00	0.30
2	crab	1	0.00	0.00
2	skunk	1	0.00	0.00
2	shark	1	0.00	0.00
3	bell	1	0.00	0.00
3	mop	1	0.00	0.30
3	dart	1	0.00	0.30
4	bow	1	0.00	0.30
4	spoon	1	0.00	0.30
4	gloves	1	0.00	0.60
5	scarf	1	0.00	0.60
5	tent	1	0.00	0.70
5	fork	1	0.00	0.90
6	swan	1	0.00	0.60
6	worm	1	0.00	0.85
6	pig	1	0.00	1.00
7	bowl	1	0.00	1.00
7	tap	1	0.00	1.04
7	knife	1	0.00	1.15
8	tie	1	0.00	1.08
8	fan	1	0.00	1.20
8	pipe	1	0.00	1.45
9	leaf	1	0.00	1.51
9	key	1	0.00	1.58
9	hand	1	0.00	2.49
10	camel	2	0.00	0.00
10	donkey	2	0.00	0.00
10	lizard	2	0.00	0.00
11	printer	2	0.00	0.00
11	teabag	2	0.00	0.00
12	croissant	2	0.00	0.00

12	lemon	2	0.00	0.00
12	onion	2	0.00	0.30
13	scissors	2	0.00	0.30
13	carrot	2	0.00	0.30
13	kettle	2	0.00	0.48
14	keyboard	2	0.00	0.48
14	plaster	2	0.00	0.48
14	orange	2	0.00	0.60
15	feather	2	0.00	0.60
15	compass	2	0.00	0.60
15	balloon	2	0.00	0.70
16	candle	2	0.00	0.70
16	apple	2	0.00	0.70
16	guitar	2	0.00	0.90
17	wheelbarrow	3	0.00	0.00
17	tambourine	3	0.00	0.00
17	strawberry	3	0.00	0.00
18	coconut	3	0.00	0.00
18	tomato	3	0.00	0.48
18	banana	3	0.00	0.70
19	whale	1	0.25	0.00
19	seal	1	0.25	0.30
19	bat	1	0.25	0.78
20	jug	1	0.25	0.30
20	mask	1	0.25	0.70
20	cheese	1	0.25	0.78
21	torch	1	0.25	0.85
21	socks	1	0.25	0.95
21	bone	1	0.25	1.11
22	stamp	1	0.25	1.20
22	tooth	1	0.25	1.20
22	belt	1	0.25	1.23
23	plate	1	0.25	1.28
23	brick	1	0.25	1.28
23	egg	1	0.25	1.76
24	waffle	2	0.25	0.00
24	igloo	2	0.25	0.00
24	lipstick	2	0.25	0.30
25	doorknob	2	1.18	N/A
25	flower	2	1.20	1.56
25	chainsaw	2	1.20	N/A
26	shaver	2	1.20	N/A
26	napkin	2	1.22	N/A
26	corkscrew	2	1.22	N/A
27	sword	1	1.21	0.70
27	pan	1	1.21	N/A

27	coin	1	1.27	0.85
28	chopsticks	2	1.26	N/A
28	canoe	2	1.28	0.00
28	grenade	2	1.29	0.00
29	phone	1	1.29	N/A
29	wreath	1	1.35	0.00
29	box	1	1.35	1.92
30	paintbrush	2	1.31	0.00
30	necklace	2	1.36	0.00
30	cracker	2	1.36	N/A
31	suitcase	2	1.37	0.30
31	light bulb	2	1.41	N/A
31	rifle	2	1.41	0.30
32	microphone	3	1.49	1.08
32	colander	3	1.50	0.00
32	toilet roll	3	1.52	N/A
33	hanger	2	1.60	N/A
33	bunk bed	2	1.62	N/A
33	dustpan	2	1.64	0.00
34	crowbar	2	1.69	N/A
34	lolly	2	1.70	N/A
34	jacket	2	1.73	0.90
35	abacus	3	1.69	0.00
35	picture frame	3	1.73	N/A
35	envelope	3	1.74	0.85
36	pill	1	1.71	N/A
36	shell	1	1.75	1.32
36	rock	1	1.78	1.45
37	teddy bear	3	1.75	0.00
37	walking stick	3	1.78	N/A
37	saxophone	3	1.83	0.00
38	trophy	2	1.89	0.00
38	post-box	2	1.91	N/A
38	anvil	2	1.98	0.00
39	hose	1	1.83	0.00
39	sledge	1	1.91	N/A
39	bread	1	1.99	1.42
40	statue	2	2.08	0.60
40	hole punch	2	2.11	N/A
40	egg timer	3	2.11	N/A
41	scraper	2	2.25	N/A
41	whisk	1	2.31	N/A
41	paddle	2	2.33	0.00
42	cable	2	2.24	1.04
42	speaker	2	2.27	1.71
42	medal	2	2.27	0.30

43	nuts	1	2.32	N/A
43	log	1	2.37	N/A
43	cake	1	2.38	0.95
44	life ring	2	2.34	N/A
44	cd	2	2.47	0.00
44	scraper	2	2.51	N/A
45	coffee grinder	4	2.38	N/A
45	fire hydrant	4	2.39	N/A
45	parking meter	4	2.52	N/A
46	gravestone	2	2.66	N/A
46	sarong	2	2.70	N/A
46	funnel	2	2.71	N/A
47	flask	1	2.60	0.30
47	weight	1	2.77	1.63
47	torch	1	2.78	0.85
48	note	1	2.79	N/A
48	vase	1	2.82	0
48	tool	1	2.93	N/A

C. Appendix: Experiment 3 - Experimental items

The introduction facts and test facts presented in Experiment 3. The first four sentences denote the introduction facts. The first two introduction facts are world knowledge facts and the second two are made up facts. The introduction fact printed in bold denotes the critical introduction fact. The fifth sentence denotes the test fact.

Correct and world knowledge test facts	
1	**The Thoroughbred breed of horses originated when English mares were crossbred with Arabian stallions.** According to his older sister, Mozart began composing pieces of music at the age of five. Toby trapped a spider in his house and let it out through the window. Cathy recently earned her brown belt in Tae Kwon Do. Thoroughbred is a cross between English mares and Arabian stallions.
2	**Mongolia has the lowest population density of any country in the world.** The Mousetrap by Agatha Christie is the longest continuously running play in the world. Beth was grounded after crashing her parents' car for the second time. Scott had to wear an arm cast for a month after falling off his bike. Mongolia has the lowest population density of any country.
3	**The Andes, running along the western coast of South America, is the world's longest terrestrial mountain range.** The World Heritage list names a site in Turkey as the archaeological site of Troy as described by Homer. Oliver bought his mother flowers and a card for her birthday. Carrie stayed in to work on her thesis instead of going to the pub with her friends. The Andes is the world's longest terrestrial mountain range
4	**The constellation Lepus represents a hare hunted by Orion and is best seen in February.** The deadliest natural disaster of any kind was the 1931 China floods, killing millions. Lucy bought a book at the airport to read on her flight home. William spent the sunny day playing Frisbee in the park with his brother. The constellation Lepus, a hare being hunted by Orion, is best seen in February

5	**The largest body of fresh water in the world is Lake Superior in North America.** Monaco is geographically the smallest UN member state. Tricia offered her seat on the bus to an elderly man who was standing. Joe was forced to retire from football when he broke his leg. Lake Superior is the largest body of fresh water in the world.

6	**The hyoid bone in the throat is the only bone in the human body not connected to another.** The Congo is the only river that flows both north and south of the equator, crossing it twice. Keith started studying in the library to prepare for his chemistry exam. Sam spent Friday afternoon playing golf with people from his work. The only bone in the human body not connected to another is the hyoid bone.

7	**Blood type AB+ is the only type in which red blood cells are not compatible with any other blood type.** The name "banana" is derived from the Arabic word "banan", meaning finger. Kevin was too tired to cook so he bought McDonald's for dinner. Edward fell asleep in biology class after staying up last night to party. Red blood cells in type AB+ blood are not compatible with any other blood type.

8	**The Marabou Stork, located in Africa, has the largest wingspan of any land bird.** The longest railway in the world is the Trans-Siberia Railway, spanning eight time zones. Geordie banged his shin again the coffee table on his way to the kitchen. Raymond played the guitar in a local pub every Friday night. Located in Africa, the Marabou Stork has the largest wingspan of all land birds.

Incorrect and world knowledge test facts

9	**The king of hearts is the only king without a moustache on a standard playing card.** The Palatinal Crypt in Budapest is the burial place of the Hungarian branch of the Habsburg dynasty. Cassie was going to go to the Bahamas for vacation before she got caught the flu. Damien took the train to London to visit with his family for Easter. The king of spades is the only king without a moustache in a set of playing cards.

10	**The Bhut Jolokia, or Ghost Chilli, is the hottest chilli pepper in the world.** In 1902, the top of the Eiffel Tower had to be reconstructed after the tower was struck by lightning." Ross is going on vacation in Cuba with some of his friends. Dianne took her dog for a walk in the park before her yoga class. The Red Savina habanero is the hottest chilli pepper in the world.
11	**The air around a lightning bolt is superheated to about five times the temperature of the Sun.** The world's smallest winged insect, the Tanzanian parasitic wasp, is smaller than the eye of a housefly. Jack went to the grocery store to buy chicken for dinner. John has recently moved to Michigan to pursue a career in music. The air around a lightning bolt is heated to about 15 times the temperature of the sun.
12	**The largest living reptile is the saltwater crocodile which inhabits parts of Asia and Australia.** While Roman Catholicism is the official state religion of Haiti, Voodoo is still practiced in rural areas. Drew played the piano in front of friends and family at his recital Alex was caught in the rain without an umbrella on her walk home from work. The siamese crocodile is the largest living reptile, inhabiting parts of Asia and Australia.
13	**The geologic era that came before the current Cenozoic is termed the Mesozoic era.** The largest movie theatre in the world is Radio City Music Hall in New York City. Shannon traveled to Italy to view the art galleries with her friends. Patrick won the silver medal in the javelin toss for his track and field team. The Paleozoic era came before the current Cenezoic era.
14	**Giraffes are the only animals born with horns, and both males and females are born with them.** Lake Pontchartrain Causeway at New Orleans in the US is the world's largest bridge. Tonia began to develop back problems from sitting at her computer too long. Nathan was asked to baby sit his little brother while his parents went to a show. Rhinos are the only animals born with horns, both males and females.
15	**Only stars that weigh 15 times more than our Sun or more will become black holes after death.** The Ojibwa are a tribe of Native Americans indigenous to the plains of North America. Claire injured her shoulder while helping her friend move furniture to a new flat. Philip moved to Aberdeen so he could take care of his parents. Only stars weighing 15 times or less than our sun become black holes after their death.

16	**In the early 16th century Michelangelo designed the Vatican's Swiss Guard uniform, which is still worn.** The Indo-Pacific Sailfish is believed to be the fastest fish in the world, and has been clocked at 68 mph. Leanne spotted a deer in her backyard when she was out gardening. Sandra took her cousin to the theme park for her birthday. Raphael designed the uniform of the Swiss guards, which they still wear today.

Correct and made up test facts

17	The flag of the Philippines is flown with the blue portion on top during peacetime and red on top in wartime. The dachshund dog breed may date back to ancient Egypt, but were bred in Europe originally to hunt badgers. **Heather scored the winning goal for her hockey team in the semi-finals.** Taylor spent all day playing football video games with his brother. In the semi-finals Heather scored the winning goal for her hockey team.

18	Augustus, not Julius Caesar, was the first emperor of Ancient Rome. The brain of Albert Einstein is currently held at McMaster University in Canada. **Spencer sat out after he sprained his ankle in his rugby game.** Samantha painted a picture of a sunset for her art exhibition. Spencer sat out of his rugby game after spraining his ankle

19	In Norse mythology, the gods were mortal and had to eat golden apples in order to not age. Aconcagua in South America is the highest mountain outside of Asia. **Tory made a dove disappear for her school's talent show.** Hannah recently went to a rock concert to see her favourite band. For her school's talent show Tory made a dove disappear

20	The longest muscle in the human body is the sartorius, located in the thigh. A Hindu temple dedicated to the rat goddess Karni Mata in India houses more than 20 000 rats. **Evan lost all his money playing five-card draw at the casino.** Mark sprained his wrist while waterskiing at his cottage. Evan lost all his money at the casino playing five card draw

21

The word safari means “journey” in Swahili.

Aluminium is the most abundant metal in the Earth’s crust.

Tommy went to his grandmother’s house so he could swim in the pool.

Elisabeth went to the store to buy a Christmas present for her brother.

Tommy went to his grandmother's house to swim in the pool.

22

The second Summer Olympic Games were held in Paris, France.

The durian is widely revered in Southeast Asia as the ìKing of Fruitsî and is known for its thorn-covered husk.

Victor purchased a songbook to help him practice playing the saxophone.

Jason went to the mall on the weekend and bought a pair of shoes.

Victor bought a songbook to help him practice the saxophone.

23

The name for a collective group of rhinos is a ìcrashî of rhinos.

Cumulonimbus clouds are vertically developed and are the clouds of rain and storms.

Jean cooked a pork roast for her family’s Christmas dinner.

Mary got a ticket for parking her car too close to a fire hydrant.

For her family's Christmas dinner Jean cooked a pork roast

24

Five different species of whale are classified as endangered, including the Blue Whale.

France was the third nation to launch a satellite into orbit with the Asterix.

Peter skipped work for the day to see his daughter’s school play.

Grace moved to Glasgow to start her new job at the university.

Peter skipped work to see his daughter's school play.

Incorrect and made up test facts

25

Riders from Belgium have won the Tour de France the second most times next to France.

The Virginia opossum is the only marsupial indigenous to North America.

Ruby booked her flight to Barcelona where she was going on her holiday.

Becky tried to get a campfire going while her friends set up the tents.

Ruby booked her flight to Venice where she was going on holiday.

26	The San Diego Zoo in California, US has the largest collection of animals in the world. Halley's Comet has one of the highest velocities relative to Earth in the solar system. **Kendrick took the bus to get to his football practice.** Gwen brought a blanket and bathing suit for her picnic at the lake. Kendrick took the bus to get to his rugby practice.

27	The Great Belt is a strait between the main Danish islands of Zealand and Fyn. Jupiter's moon Ganymede is the largest moon in the solar system, and is larger than the planet Mercury. **Lily recently had her job interview at the Italian restaurant near her home.** Lauren is travelling to Africa to volunteer with children for a month. Lily recently had a job interview at a Thai restaurant near her home.

28	The onager was an ancient Roman siege weapon which derived its name from its kicking action, like that of a donkey. The stratum corneum is a thick layer of skin that covers human finger tips and soles of feet. **Daniel took a year off school to travel around South America.** Karen went to the cinema with her sister and bought a chocolate bar. Daniel took a year off school to travel around Europe.

29	The anemometer is an instrument that measures the pressure, force, and velocity of wind. The Ancient Egyptians used mastaba tombs to celebrate their deified pharaohs after their death. **Marie moved to Vancouver where she would be attending university.** Sean lost all his phone numbers when he left his mobile phone at the train station. Marie moved to California where she would attend university.

30	The deepest lake in the world is Lake Baikal in Siberia. English explorer James Cook and his crew were the first to encounter Australia's eastern coastline. **Kim recently moved to Japan to teach English to schoolchildren.** Jacob ironed his shirt to prepare for the graduation ball. Kim recently moved to China to teach English to schoolchildren.

31	The Algonquin believed the Northern Lights to be their ancestors dancing around a ceremonial fire. In Greek mythology the first labor of Heracles was to kill the Nemean Lion. **Stephanie left work early on Monday to go to her dentist appointment.** Neil went for a hike through the mountains on his trip in New Zealand. On Monday Stephanie left work early to go see a chiropractor.

32	Knapping refers to creating a tool with a cutting edge or blade out of stone, such as flint. Of the Seven Wonders of the Ancient World, the Lighthouse of Alexandria was the tallest. **Anna ordered lobster at the restaurant for her birthday dinner.** Gus is moving to Ireland to start his new job as a math teacher. For her birthday dinner at the restaurant Anna ordered steak.

D. Appendix: Experiments 4 & 5 - Experimental Items

The introduction facts and test facts presented in Experiment 4 and 5. The first four sentences denote the introduction facts. The introduction fact printed in bold denotes the critical introduction fact. The fifth sentence denotes the correct test fact and the sixth sentence denotes the incorrect test fact.

Easy test facts

1	**Popeye is the name of the comic strip character who eats spinach to increase his strength!** Monaco located in Europe is geographically the smallest UN member state. The name “banana” is derived from the Arabic word “banan”, meaning finger. A chameleon is a lizard that changes its colour to match the surroundings. Popeye is a comic strip character and eats a lot of spinach. Popeye is a comic character and eats a lot of carrots.
2	**The zebra is a horse-like animal with black and white stripes that mostly live in Africa.** The top of the Eiffel Tower had to be reconstructed after the tower was struck by lightning. "Augustus, not Julius Caesar, was the first emperor of Ancient Rome." A surgeon is a medical doctor who specializes in cutting the body. The zebra is an animal which has black and white stripes. The zebra is an animal which has black and white checked patterns.
3	**Paris is the capital of France, situated on the river Seine in the northern part of the country.** France was the third nation to launch a satellite into orbit with the Asterix. Riders from Belgium have won the Tour de France the second most times next to France. A cyclops is a legendary one-eyed giant in Greek mythology. Paris is the capital of France and is located in the northern part of the country. Nancy is the capital of France and is located in the north of the country.

Lava is molten rock that runs down the side of a volcano during an eruption.
The Ancient Egyptians used mastaba tombs to celebrate their pharaohs after their death.
The Algonquin believed the Northern Lights to be their ancestors dancing around a fire.
4 A glider is an airplane without an engine.

Lava is molten rock that runs down the side of a volcano.
Lava is molten rock that runs down the side of big trees.

Tennis is the sport associated with Wimbledon and is played with a ball and a racket.
Aluminium is the most abundant metal in the Earth’s crust.
The largest movie theatre in the world is Radio City Music Hall in New York City.
5 Jupiter is the largest planet in the solar system and the 5th planet from the sun.

Tennis is very much associated with Wimbledon in south England.
Golf is very much associated with Wimbledon in south England.

Fossils are the remains of plants and animals that are usually found in stone.
The Andes in South America is the world’s longest terrestrial mountain range.
"The San Diego Zoo in California, US has the largest collection of animals in the world."
6 A cobra is the type of snake Asian snake charmers use.

Fossils are the remains of plants and animals and are found in stone.
Fossils are the remains of plants and animals and are found in rivers.

A puck is a rubber object that is hit back and forth by hockey players on an ice rink.
In Norse mythology, the gods were mortal and had to eat golden apples in order to not age.
The brain of Albert Einstein is currently held at McMaster University in Canada.
7 Meteors are astronomical bodies that enter the earth's atmosphere.

A puck is an object that is hit by hockey players on an ice rink.
A ball is an object that is hit by hockey players on an ice rink.

Einstein is the last name of the man who proposed the theory of relativity in 1905.
James Cook and his crew were the first to encounter Australia’s eastern coastline.
Five different species of whale are classified as endangered, including the Blue Whale
8 The pacific is the largest ocean in the world and located between Asia and America.

Einstein was the man who first proposed the theory of relativity.
Newton was the man who first proposed the theory of relativity.

A migraine is a severe headache that returns periodically and often is accompanied by nausea.
The world’s smallest winged insect is the Tanzanian parasitic wasp.
Cumulonimbus clouds are vertically developed and are the clouds of rain and storms.
9 Prune is the name of a dried plum.

A migraine is a headache which often is accompanied by nausea.
A migraine is a headache which often is accompanied by hunger.

Presley is the last name of the singer who recorded “Heartbreak Hotel” and “All shook up”.
"Jupiter’s moon Ganymede is the largest moon in the solar system, and is larger than Mercury."
The Mousetrap by Agatha Christie is the longest continuously running play in the world.
10 Nomads are desert people who wander instead of living in one place.

Presley is the singer who recorded “Heartbreak Hotel” and “All shook up”.
Johnny Cash is the singer who recorded “Heartbreak Hotel” and “All shook up”.

Shakespeare is the last name of the author who wrote the tragedy “Romeo and Juliet”.
The stratum corneum is a thick layer of skin that covers human finger tips and soles of feet.
The Palatinal Crypt is the burial place of the Hungarian branch of the Habsburg dynasty.
11 Insomnia is the word meaning an inability to sleep.

Shakespeare is the author of the tragedy "Romeo and Juliet".
Hemingway is the author of the tragedy “Romeo and Juliet”.

Photosynthesis is the process by which plants convert light energy into chemical energy.
The anemometer is an instrument that measures the pressure, force, and velocity of wind.
Aconcagua in South America is the highest mountain outside of Asia.
12 A tornado is a cyclone that occurs over land and can destroy whole houses.

Photosynthesis is the process by which plants produce energy.
Erosion is the process by which plants produce energy.

Titanic is the name of the supposedly unsinkable ship that sunk on its maiden voyage.
The deadliest natural disaster of any kind was the 1931 China floods, killing millions.
The Great Belt is a strait between the main Danish islands of Zealand and Fyn.
13 Gold is associated with a 50th wedding anniversary.

Titanic is the supposedly unsinkable ship that sunk on its first voyage.
Titanic is the supposedly unsinkable ship that sunk on its third voyage.

	Slalom is the type of ski race in which a downhill skier makes sharp turns around poles.
	In Greek mythology the first labour of Heracles was to kill the Nemean Lion.
	According to his older sister, Mozart began composing pieces of music at the age of five.
14	Bogart is the last name of the male star of the movie ""Casablanca"" filmed in the 19th century.
	Slalom is a ski race in which the skier makes turns around poles.
	Slalom is a ski race in which the skier makes turns around cubes.

	The Sahara is the largest hot desert on earth located in the north of Africa.
	The deepest lake in the world is Lake Baikal in Siberia.
	The second Summer Olympic Games were held in Paris, France.
15	A birdie is a term in golf referring to a score of one under par on a particular hole.
	The Sahara is the largest hot desert located in the north of Africa.
	The Sahara is the largest hot desert located in the north of Europe.

	Atlantis is the name of the island-city believed since antiquity to have sunk into the ocean.
	The longest muscle in the human body is the sartorius, located in the thigh.
	The Indo-Pacific Sailfish is believed to be the fastest fish in the world.
16	Raisins are dried grapes.
	Atlantis is believed since antiquity to have sunk into the ocean.
	Pompey is believed since antiquity to have sunk into the ocean.

Difficult test facts

	Roger Bannister is the name of the man who ran the mile in under 4 minutes in 1954.
	Decibel is the unit of sound intensity.
	Tidal waves are giant ocean waves caused by earthquakes.
17	The Ojibwa are a tribe of Native Americans indigenous to the plains of North America.
	Roger Bannister first ran the mile in under 4 minutes in 1954.
	Roger Bannister first ran the mile in under 4 minutes in 1974.

	Mongolia has the lowest population density of any country in the whole world.
	Sistine is the name of the chapel in which Michelangelo painted the ceiling.
	A dermatologist is a medial doctor who specializes in diseases of the skin.
18	The Virginia opossum is the only marsupial indigenous to North America.
	Mongolia has the lowest population density of any country.
	Canada has the lowest population density of any country.

Florence is the city where the original of Michelangelo's statue of David is located.

Balsa is the name of the lightest wood known and is native to South America.

Armstrong is the name of the first person to set foot on the moon in 1969.

19 The Congo is the only river that flows both north and south of the equator, crossing it twice.

Florence is the city where the statue of David is located.

Rome is the city where the statue of David is located.

The Lepus constellation represents a hare hunted by Orion and is best seen in February.

Checkers is the game in which men are crowned.

Moscow is the capital of Russia located in the European part of the country.

20 The durian is widely revered in Southeast Asia as the “King of Fruits”.

The Lepus constellation represents a hare being hunted by Orion.

The constellation Lepus represents a hare being hunted by Taurus.

Lake Superior in North America is the largest body of fresh water in the world.

Venice located in northern Italy is an Italian city known for its canals.

Plasma is the liquid portion of blood, in which the blood cells are suspended.

21 Only stars that weigh 15 times more than our Sun or more will become black holes after death.

Lake Superior is the largest body of fresh water in the world.

Lake Superior is the largest body of fresh water in South Africa.

The hyoid bone in the throat is the only bone in the human body not connected to another.

Blubber is a thick layer of fat on a whale and is used to manufacture cosmetics.

Venison is the name of deer meat.

22 The air around a lightning bolt is superheated to about five times the temperature of the Sun.

The hyoid bone is the only human bone not connected to another.

The jar bone is the only human bone not connected to another.

AB+ is the only blood type in which red blood cells are not compatible with any other type.

A strike is the term for hitting a volleyball hard into the opponent's court.

Siamese cats have blue eyes and are typically long lived.

23 The name for a collective group of rhinos is a “crash” of rhinos.

AB+ blood type is not compatible with any other blood type.

A+ blood type is not compatible with any other blood type.

The Marabou Stork, located in Africa, has the largest wingspan of any land bird.
A lamb is a young sheep.
An ostrich is a bird that cannot fly and is the largest bird on earth.
24 "Knapping" refers to creating a tool with a cutting edge or blade out of stone, such as flint.

The Marabou Stork has the largest wingspan of all land birds.
The Marabou Stork has the largest wingspan of all sea birds.

The king of hearts is the only king without a moustache on a standard playing card.
Hibachi is a small Japanese stove used for outdoor cooking.
A kilt is a short pleated skirt worn by men in Scotland.
25 Halley's Comet has one of the highest velocities relative to Earth in the solar system.

The king of hearts is the only king without a moustache.
The king of hearts is the only king without a crown.

The Bhut Jolokia, or Ghost Chilli, is the hottest chilli pepper that can be found on earth.
Treason is a crime in which a person purposely betrays his country.
Tarantula is the name of the large hairy spider that lives near bananas.
26 Of the Seven Wonders of the Ancient World, the Lighthouse of Alexandria was the tallest.

The Bhut Jolokia is the hottest chilli pepper in the world.
The Red Savina is the hottest chilli pepper in the world.

Post is the last name of the first flier to fly solo around the world in 1931.
An odometer is an automobile instrument that measures mileage.
Hibernation is the name of the long sleep some animals go though during the entire winter.
27 A Hindu temple dedicated to the rat goddess Karni Mata houses more than 20 000 rats.

Post is the first man who flew solo around the world in 1931.
Post is the first man who flew solo around the world in 1961.

The saltwater crocodile, which inhabits parts of Asia and Australia, is the largest living reptile.
Arson is the name of the crime in which a building or property is purposely set on fire.
The Indian ocean is located between Africa and Australia.
28 Lake Pontchartrain Causeway at New Orleans in the US is the world's largest bridge.

The saltwater crocodile living in Asia and Australia is the largest reptile.
The Siamese crocodile living in Asia and Australia is the largest reptile.

The Mesozoic era is the geologic era that came before the current Cenozoic.

Mercury is the only liquid metal at room temperature and is used in thermometers.

Jane is the name of Tarzan's girlfriend who met him in the jungle.

29 While Catholicism is the official state religion of Haiti, Voodoo is still practiced in rural areas.

The Mesozoic era came before the current Cenozoic era.

The Paleozoic era came before the current Cenozoic era.

Giraffes are the only animals born with horns, and both males and females are born with them.

Lindberg was the first person to complete a solo flight across the Atlantic ocean.

A flush is a poker hand in which all of the cards are of the same suit.

30 The World Heritage list names a site in Turkey as the archaeological site of Troy.

Giraffes are the only animals born with horns, both males and females.

Rhinos are the only animals born with horns, both males and females.

Rhine is the name of the river on which the German city Bonn is located.

Hook is the last name of the villainous captain in the story “Peter Pan”.

Cleopatra is the name of the Egyptian queen who joined forces with Mark Antony of Rome.

31 The dachshund dog breed may date back to ancient Egypt, but were bred in Europe.

Rhine is the river on which the German city Bonn is located.

Rhine is the river on which the German city Hamburg is located.

Michelangelo designed the Vatican’s Swiss Guard uniform In the early 16th century.

Japan located in Asia is the country for which the yen is the monetary unit.

Knot is the word that means a nautical mile per hour and is used to measure speed.

32 The longest railway in the world is the Trans-Siberia Railway, spanning eight time zones.

Michelangelo designed the uniform of the Swiss guards in the 16th century.

Michelangelo designed the uniform of the Swiss guards in the 18th century.

E. Appendix: Experiment 5 – Mathematical Equations

Correct mathematical problems	Incorrect mathematical problems
366 – 119 = 247	764 – 367 = 297
456 – 79 = 377	933 – 477 = 356
541 – 38 = 503	121 – 65 = 46
802 – 25 = 777	254 – 68 = 176
743 – 145 = 598	962 – 96 = 766
692 – 295 = 397	746 – 239 = 407
910 – 81 = 829	75 – 37 = 48
733 – 78 = 655	113 – 56 = 47
104 – 79 = 25	143 – 26 = 127
811 – 372 = 439	185 – 127 = 68
116 – 79 = 37	612 – 84 = 538
384 – 269 = 115	123 – 25 = 88
712 – 598 = 114	451 – 43 = 418
301 – 15 = 286	787 – 249 = 548
151 – 14 = 137	895 – 38 = 858
655 – 69 = 586	246 – 78 = 158

F. Appendix: Experiment 6 – Garden Path Sentences

Experimental items

1	The new city council argued the position of the young radical mayor was immoral.
2	The ambassador wrote the article which he used impressed an old friend of his.
3	The detective saw the man with a gun in his hand fall over.
4	I suppose that the clerk knows the woman wearing that outrageous peacock hat is crazy.
5	The criminal confessed his sins which upset kids harmed too many people.
6	Before the king rides his horse which is white and beautiful is always groomed.
7	Wherever Alice walks her sheep dog which is shaggy will follow.
8	Though Hilda finally agreed to sing the song she chose turned out to be just awful.
9	While Mary was mending the grandfather clock in the hall started to chime.
10	As the carpenter builds the table that sloped breaks in the middle.
11	When the scientist was teaching the students of botany took the books away.
12	Whilst the janitor polished the floor of the hall shone rather brightly.
13	When the cleaners were rubbing the paint that dried stained the new carpet.
14	Because Paul's son likes to visit people who are older think he's a terrific neighbour.
15	His second wife will claim the family inheritance which was very large belongs to her.
16	The secretary failed to mention the accounting error which was very large was her own fault.
17	This morning, Sam remembered his exams to which he wasn't looking forward to would be soon.

18 The travel agent confirms the reservation of his received a stamp of approval from the boss.

19 The journalist wrote the book from upstairs caused very great interest.

20 The speaker concluded his lecture which was interesting but technical had been a success.

21 The teacher believed the pupil in the cafeteria failed to meet the requirements.

22 Since Jay always jogs a mile and a half really seems like a very short distance to him.

23 The manager taught the employee that he chose required a better filing system.

24 The lecturer wrote the speech that upset people annoyed all of the politicians.

25 The visitor reads the adverts from Ireland encouraged the young people to smoke.

26 Yesterday, Sally found out the answer to the difficult physics problem was in the book.

27 The head teacher declared the holiday that he liked seemed to cheer up the staff.

28 After the judge decided the verdict of the trial caught the old man's attention.

29 The government cautioned the companies that lied promoted a big increase in drinking.

30 The cabinet minister proposed the mining policy for the new region would be a great benefit.

31 As the woman edited the magazine about fishing amused all the reporters.

32 After Mary drank the water which looked strange was discovered to be polluted.

33 Last week, Tom heard the latest gossip about the new neighbours wasn't true.

34 The historian proves the theory from Oxford resolved a very messy dispute.

35 As the artist paints the picture of the roses pleases all the critics greatly.

36 When the ambassador negotiated the treaty about arms upset many of the civilians.

37	Though George kept on reading that book about a science fiction story really bothered him.
38	While the motorist parks the lorry that was noisy rushes along the high street.
39	As the zoologists were breeding the tigers from India chewed on a piece of meat.
40	As the cowboys roped the horses that had escaped charged across the ranch.

Questions

		Answer
1	Did the city council argue about the morality of the mayor's position?	Yes
2	Did the ambassador use an article?	Yes
3	Did the detective see someone fall over?	Yes
4	Is the woman crazy?	Yes
5	Did the criminal's sins harm someone?	Yes
6	Is the king's horse often groomed?	Yes
7	Does the shaggy dog follow Alice all the time?	Yes
8	Was the song awful?	Yes
9	Could Mary have mended something else than the grandfather clock?	Yes
10	Could the table have been built by someone else than the carpenter?	Yes
11	Did the students take the books away?	Yes
12	Could the janitor have polished something else than the floor?	Yes
13	Could someone else than the cleaners have stained the new carpet?	Yes

14	Do people think Paul's son is terrific?	Yes
15	Was the family inheritance very large?	Yes
16	Was the error the fault of the secretary?	Yes
17	Did Sam think about an upcoming exam?	Yes
18	Does the boss approve of the reservation?	Yes
19	Did the journalist write that the book got much attention?	Yes
20	Does the sentence say that the speaker concluded his lecture?	Yes
21	Does the sentence say that the teacher believed the pupil?	No
22	Does the sentence say that Jay always jog a mile and a half?	No
23	Does the sentence say that the manager taught the employee?	No
24	Did the lecturer necessarily write the speech?	No
25	Does the sentence say that the visitor reads adverts?	No
26	Did Sally think about the solution of a problem all by herself?	No
27	Does the head teacher plan to go on holiday?	No
28	Does the sentence say that the judge decided the verdict?	No
29	Did the government caution the companies?	No
30	Does the sentence say that the cabinet minister proposed a policy of mining?	No
31	Did the woman amuse all the reporters?	No
32	Does the sentence say that Mary drank the strange looking water?	No

33	Does the sentence say that Tom heard some gossip last week?	No
34	Did the historian necessarily prove the theory?	No
35	Did the artist necessarily paint a picture of roses?	No
36	Does the sentence say that the ambassador negotiated a treaty?	No
37	Did George find the story really funny?	No
38	Did the motorist park the lorry?	No
39	Does the sentence say that the zoologists bred the tigers?	No
40	Does the sentence say that the cowboys roped the horses?	No

G. Appendix: Masked characters

Original letter	Masked letter	Original letter	Masked letter	Original letter	Masked letter
A	B	V	C	q	p
B	C	W	M	r	c
C	D	X	S	s	z
D	F	Y	T	t	l
E	G	Z	V	u	s
F	H	a	n	v	x
G	K	b	h	w	m
H	X	c	s	x	n
I	t	d	b	y	q
J	P	e	c	z	v
K	Q	f	t	1	I
L	R	g	q	2	S
M	W	h	k	3	F
N	Y	i	l	4	A
O	P	j	f	5	H
P	G	k	h	6	C
Q	B	l	t	7	J
R	P	m	w	8	K
S	N	n	c	9	L
T	F	o	v	0	U
U	D	p	g		

H. Appendix: Experiment 7 – Garden Path Sentences

Experimental Items

1	The city council argued the position of the radical mayor insinuates a bad immoral outlook.
2	The young ambassador wrote the article which he once used impressed a long lost friend.
3	The hired detective saw the man with a large gun humiliate the scared shop assistant.
4	I suppose he knows the woman wearing that outrageous hat performed with quite crazy stunts.
5	The known criminal confessed his sins which upset many kids disturbed too many kind people.
6	Before the king rides his horse which is very beautiful experiences lots of good care.
7	Wherever Alice enjoys walking her dog which is quite shaggy remembers his usual pet treats.
8	Though Hilda agreed singing the song she had quickly chosen completed in a disappointing way.
9	While Mary was mending the clock in the dark hall frightened her with its chime.
10	As the carpenter builds the table that bends and slopes collapses abruptly in the middle.
11	When the scientist taught the students of biology and chemistry surrounded all the lab equipment.
12	Whilst the janitor polished the floor of the school hall metamorphosized into a beautiful sight.
13	When the cleaners rubbed the paint that ran and dried destroyed the brand new carpet.
14	Because James likes visiting the people who are much older described him with great pleasure.
15	His second wife claims the inheritance which was very large disappeared from his hidden safe.
16	The secretary missed mentioning the error which caused the loss initiated in her own office.
17	This morning, Sam remembered his exams that he was dreading calculated his final year mark.

18 The travel agent confirms the reservation of the young man succeeded over the other bookings.

19 As the journalist wrote the book from the room upstairs increased interest amongst the readers.

20 The speaker happily concluded his lecture which he struggled with relinquished his previous bad reputation.

21 The teacher strongly believed the pupil in the cafeteria abandoned the entire school rules.

22 Since Jay always jogs a mile from school to home definitely seems short to him.

23 The office manager taught the employee that he last chose dominated all the other staff.

24 The old lecturer wrote the speech that upset most people disrupted all of the politicians.

25 The visitor reads the adverts in the shop window encouraged the people to smoke.

26 Last night, Sally discovered the answer to the difficult problem contradicted her friend's first attempt.

27 The head teacher declared the holiday that he really liked retrieved the old staff morale.

28 After the judge decided the verdict of the long trial captivated the old man's attention.

29 The government memo cautioned the companies that had often lied advertised alcohol to vulnerable youths.

30 The cabinet minister proposed the policy for the new region incorporates great benefits for all.

31 As the woman edited the magazine about fishing in Britain delighted all the reporters.

32 After Mary had drank the water which looked strange evaporated into a purple cloud.

33 Last week, Tom heard the gossip about the new neighbours possesses no truth.

34 The wise historian proves the theory from an Oxford professor reiterates ideas from the past.

35 As the artist paints the picture of the night scene astonishes all the critics greatly.

36 When the ambassador negotiated the treaty about the past events infuriated many of the civilians.

37	Though George continued reading that book about a fictional story extremely bothered him last night.
38	While the motorist parks the lorry that was so noisy whirlwinds along the high street.
39	Whilst zoologists were feeding the tigers from South East India struggled to chew their food.
40	As the cowboys roped the horses that had escaped continued galloping across the ranch.

Questions

		Answer
1	Did the city council argue about the morality of the mayor's position?	Yes
2	Does the sentence say that the ambassador used an article?	Yes
3	Did the man with the gun humiliate someone?	Yes
4	Does the woman do something crazy?	Yes
5	Did the criminal's sins disturbed someone?	Yes
6	Does the sentence say that the king's horse got some care?	Yes
7	Does the shaggy dog remember his pet treats?	Yes
8	Was the song's ending disappointing?	Yes
9	Could Mary have mended something else than the clock?	Yes
10	Could the table have been built by someone else than the carpenter?	Yes
11	Does the sentence say that the students surrounded the lab equipment?	Yes
12	Could the janitor have polished something else than the floor?	Yes
13	Could someone else than the cleaners have destroyed the new carpet?	Yes

14	Do people think that James is kind?	Yes
15	Does the sentence say that the family inheritance disappeared?	Yes
16	Was the error the fault of the secretary?	Yes
17	Did Sam think about an exam?	Yes
18	Was the reservation successful?	Yes
19	Does the sentence say that the readers were interested in the book from upstairs?	Yes
20	Does Jay think that a mile is not very much?	Yes
21	Does the sentence say that the speaker concluded his talk?	No
22	Does the sentence say that the teacher believed the pupil?	No
23	Does the sentence say that the manager taught the employee?	No
24	Did the lecturer necessarily write the speech?	No
25	Does the sentence say that the visitor reads adverts?	No
26	Did Tim think about the solution of a problem all by himself?	No
27	Does the head teacher plan to go on holiday?	No
28	Does the sentence say that the judge decided the verdict?	No
29	Did the government caution the companies?	No
30	Does the sentence say that the cabinet minister proposed a policy?	No
31	Did the woman delighted all the reporters?	No
32	Does the sentence say that Mary drank the water?	No

33	Does the sentence say that Tom heard some gossip last week?	No
34	Did the historian necessarily prove the theory?	No
35	Did the artist necessarily paint a picture of a night scene?	No
36	Does the sentence say that the ambassador negotiated a treaty?	No
37	Did George find the story really funny?	No
38	Did the motorist park the lorry?	No
39	Does the sentence say that the zoologists fed the tigers?	No
40	Does the sentence say that the cowboys roped the horses?	No